



# UFBA

UNIVERSIDADE FEDERAL DA BAHIA  
ESCOLA POLITÉCNICA  
PROGRAMA DE PÓS GRADUAÇÃO EM  
ENGENHARIA INDUSTRIAL - PEI

DOUTORADO EM ENGENHARIA  
INDUSTRIAL

CLEBER NAUBER DOS SANTOS

Plataforma Multidimensional *Online* para Canais de Pesquisa,  
Desenvolvimento Tecnológico e Serviços em Instituições  
Científicas e Tecnológicas Baseada em *Business Intelligence*

SALVADOR



2023



UNIVERSIDADE FEDERAL DA BAHIA  
ESCOLA POLITÉCNICA  
PROGRAMA DE PÓS-GRADUAÇÃO EM  
ENGENHARIA INDUSTRIAL



**CLEBER NAUBER DOS SANTOS**

**Plataforma Multidimensional *Online* para Canais de Pesquisa, Desenvolvimento  
Tecnológico e Serviços em Instituições Científicas e Tecnológicas Baseada em  
*Business Intelligence***

**SALVADOR  
2023**

**CLEBER NAUBER DOS SANTOS**

**Plataforma Multidimensional *Online* para Canais de Pesquisa, Desenvolvimento Tecnológico e Serviços em Instituições Científicas e Tecnológicas Baseada em *Business Intelligence***

**Tese apresentada ao Programa de Pós-Graduação em Engenharia Industrial – PEI, Escola Politécnica, Universidade Federal da Bahia, como requisito parcial para obtenção do grau de Doutor em Engenharia Industrial.**

**Orientador: Prof. Dr. Marcelo Embiruçu de Souza  
Coorientadores: Prof. Dr. Carlos Henrique Almeida Alves e Prof. Dr. Marcio Luis Ferreira Nascimento**

**SALVADOR  
2023**

Modelo de ficha catalográfica fornecido pelo Sistema Universitário de Bibliotecas da UFBA para ser confeccionada pelo autor

---

NAUBER, Cleber

Plataforma Multidimensional Online para Canais de Pesquisa, Desenvolvimento Tecnológico e Serviços (PD&S) em Instituições Científicas e Tecnológicas (ITCs) Baseada em Business Intelligence (BI)/ CLEBER NAUBER DOS SANTOS. - Salvador, 2023.

348 f. :il color.

Orientadores: Prof. Marcelo Embiruçu de Souza.

Coorientadores: Prof. Carlos Henrique Almeida Alves.

Prof. Marcio Luis Ferreira Nascimento.

Tese (Doutorado - Engenharia Industrial) - Universidade Federal da Bahia, PEI, 2023.

1. *Business Intelligence*. 2. *Dashboard e Data Warehouse*. 3. Tripla Hélice de Inovação. 4. Canais de Pesquisa, Desenvolvimento Tecnológico e Serviços (PD&S). 5. Instituição ou Rede Federal de Educação Profissional, Científica e Tecnológica (RFEPECT) ou Universidades ou ICT. I. Embiruçu, Marcelo. II. Nascimento, Marcio Luis Ferreira. III. Alves, Carlos Henrique Almeida. IV. Título.

---

**Plataforma Multidimensional *Online* para Canais de Pesquisa, Desenvolvimento Tecnológico e Serviços em Instituições Científicas e Tecnológicas Baseada em *Business Intelligence***

**Cleber Nauber dos Santos**


Tese submetida ao corpo docente do programa de pós-graduação em Engenharia Industrial da Universidade Federal da Bahia como parte dos requisitos necessários para a obtenção do grau de doutor em Engenharia Industrial.

Examinada por:



Prof. Dr. Angelo Conrado Loula

Doutor em Engenharia Elétrica pela Universidade Estadual de Campinas, Brasil, 2011



Prof. Dr. Gesil Sampaio Amarante Segundo

Doutor em Física pela Universidade de São Paulo, Brasil, 2000



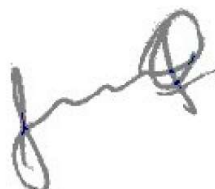
Prof. Dr. Luis Paulo Leopoldo Mercado

Doutor em Educação pela Pontifícia Universidade Católica de São Paulo, Brasil, 1998



Prof. Dr. Marcio Luis Ferreira Nascimento

Doutor em Ciência e Engenharia dos Materiais pela Universidade Federal de São Carlos, Brasil, 2004



Prof. Dr. Josealdo Tonholo

Doutor em Físico-Química pela Universidade de São Paulo, Brasil, 1997

Salvador, BA – BRASIL  
2023

Dedico este trabalho a minha família, em especial, aos meus pais e aos meus mestres e professores que, em seu tempo e de sua forma, contribuíram para minha formação como ser social, profissional, docente e pesquisador.

## **Agradecimentos**

Em primeiro lugar, a Deus, a todos os Santos, encantos e Axé; a todas as divindades e energias do universo, tanto do sagrado quanto do profano, expresso minha mais profunda gratidão e orgulho pela oportunidade e bênção de ser baiano. Que todas as tradições, crenças e caminhos espirituais se unam em uma completa harmonia e alegria, refletindo o rico sincretismo de fé e cultura na qual fui forjado.

Aos meus avós paternos e maternos, que semearam as raízes da nossa família, que me influenciaram por meio de suas histórias e experiências. Seus múltiplos ensinamentos foram inspiradores para que eu pudesse me tornar o primeiro indivíduo, em nossa rica árvore genealógica, a alcançar o mais alto nível de formação e conhecimento acadêmico.

À minha amada esposa, Suzana, e ao nosso precioso filho, Miguel, agradeço por serem minha constante fonte de inspiração e força. Vocês são a razão pela qual eu busco ser melhor a cada dia.

Aos meus pais, Guimária e Edmilton, por sua inabalável fé em mim e pelo amor incondicional que sempre me proporcionaram. Reconheço formalmente o amor que vocês me deram e dedicaram, cada um do seu jeito e estilo. O amor de vocês transcende palavras e gestos, porém tornam-se a bússola que sempre me direcionou na vida: “Vá estudar!”. Em cada passo que dou, carrego a essência da dedicação e do amor de vocês. Por mais que eu agradeça, reconheço que qualquer tentativa de expressar minha gratidão ainda será insuficiente diante da magnitude do que fizeram e fazem por mim, pelos meus irmãos e pela nossa família.

Aos meus irmãos, Macio e Luan, por estarem sempre ao meu lado, apoiando-me em cada etapa desta jornada. Em vocês busquei a inspiração e força, tendo como espelho suas proatividades, espírito de liderança e coragem para irem em busca de seus sonhos.

Aos meus sogros, Conceição e Vitorino, e aos meus cunhados Flávio e Ana, que se tornaram meus pais e irmãos de coração. Dedico a vocês o meu mais sincero e profundo agradecimento e serei eternamente grato por fazerem parte da minha vida e da construção dessa trajetória.

Ao meu orientador, Prof<sup>o</sup>. Dr. Marcelo Embiruçu, o meu respeito e admiração por você, por ter aceitado me orientar nessa longa, mas produtiva jornada, e pelas intervenções pontuais que me fizeram rever caminhos e encontrar soluções. Que esta simples mensagem possa transmitir uma fração do meu eterno apreço.

Aos meus coorientadores, Prof. Marcio Nascimento e Carlos Henrique, não posso começar sem expressar o quão profundo é o meu respeito e admiração por vocês. Em cada discussão, em cada tropeço e em cada triunfo, vocês estiveram lá, cada um dentro de suas possibilidades, não apenas como coorientadores acadêmicos, mas como um farol iluminando o caminho da minha jornada intelectual. A sabedoria de vocês transcende o conhecimento acadêmico; é temperada pela humanidade, pela paciência e por uma genuína preocupação com o meu crescimento enquanto pesquisador. Não foi apenas a ciência ou a pesquisa que vocês me ensinaram, mas também a arte de pensar, de questionar e de persistir. O título de "coorientador" não é suficiente para capturar a magnitude da sua influência nesta minha trajetória formativa. A vocês, toda minha gratidão eterna, pois passaram a fazer parte da mais seleta lista que considero terem sido meus mestres do conhecimento.

Ao IFAL e, particularmente, ao Curso de Design de Interiores, minha gratidão estendida aos meus colegas de trabalho. A colaboração de vocês foi fundamental para que pudesse concluir essa formação, uma vez que a realizei sem afastamento. Obrigado por me inspirarem institucionalmente nos momentos de dificuldades durante essa trajetória. Aqui, celebra-se não apenas o conhecimento adquirido, mas as inúmeras histórias e conexões criadas. Nosso caminho no mundo do *Design* é repleto de paixão e solução de problemas, e por isso, é com profundo reconhecimento que afirmo: obrigado por fazerem parte dessa jornada.

Aos meus mestres que moldaram minha trajetória acadêmica: Prof<sup>ª</sup>. Zélia, Prof<sup>º</sup>. Jorge Resende Couto, Prof<sup>º</sup>. Alexandre Wollner, Prof<sup>º</sup>. Cleomar Rocha e Prof<sup>º</sup>. Luís Paulo Leopoldo. Seus ensinamentos e orientações foram fundamentais para atravessar a Odisseia.

Aos meus amigos diretamente envolvidos: Carlos Henrique, Victor Sgarbi, Marcio Luis, Marcelo Embiruçu, Edson Camilo, Ciro Bezerra, Thiago Marques, Manel Moura, Matheus Cristian, Erick Araújo, dentre outros, agradeço pelas colaborações, discussões, incentivos, críticas e aprendizados compartilhados, a amizade e apoio essencial alcançar este marco.

Finalmente, a todos os demais familiares e amigos que, em algum momento da minha existência, compartilharam este objetivo, minha sincera gratidão. Dedico este trabalho a todos e todas que estiveram comigo nesta jornada. Gratidão a todos vocês.



Meu caminho pelo mundo eu mesmo traço, afinal a Bahia já me deu a régua e compasso (Gil, 1969). A percepção do desconhecido é a mais fascinante das experiências. O homem que não tem os olhos abertos para o misterioso passará pela vida sem nada ver. Tenha em mente que tudo que você aprende na escola é trabalho de muitas gerações. Receba essa herança, honre-a, acrescente a ela e, um dia, finalmente deposite-a nas mãos de seus filhos. (Autor Desconhecido)

## RESUMO

Dada a crescente importância do atendimento às demandas institucionais, produtivas e governamentais no ecossistema de Ciência, Tecnologia e Inovação, esta tese buscou resolver um problema crítico: como organizar e aprimorar a gestão dos dados de pesquisa, desenvolvimento tecnológico e prestação de serviços (PD&S) nas Instituições de Ciência e Tecnologia (ICTs), de modo a responder de maneira tempestiva e confiável a essas exigências. Para isso, o trabalho em questão apresentou uma solução tecnológica online desenvolvida para integrar e sistematizar dados públicos deste ecossistema. Isso foi registrado junto ao INPI (BR5120230010197), fundamentado em teorias de inovação, análise de negócios e visualização de dados. Utilizou como abordagem quantitativa a incorporação de técnicas de *Business Intelligence*, *data warehouse* e *dashboards* interativos. Para a validação desta tese, utilizou-se os dados da Rede Federal de Educação Profissional, Científica e Tecnológica como estudo de caso. O desafio foi superar os obstáculos da integração dos dados referentes a gestão da inovação, o que tem dificultado a tomada de decisões assertivas por parte da sociedade em geral como dos gestores das ICTs, pesquisadores e setores produtivos. Ao conceber e implementar tais soluções tecnológicas, ancoradas nas teorias de Análise de Negócios, Visualização de Dados e da teoria da Tripla Hélice de Inovação foi possível apresentar e compreender melhor os problemas a partir da disponibilidade dos dados à sociedade de forma oportuna e segura. Na metodologia, utilizou-se de uma abordagem quantitativa e empírico-dedutiva que incluiu revisão de literatura, consulta a usuários, identificação de fontes de dados e validação do sistema por nove especialistas. Os resultados indicaram que a ferramenta foi essencial para consolidar informações de múltiplas fontes, auxiliando a sociedade em geral na resposta rápida e confiável às demandas apresentadas. A solução proposta também apresentou uma interface amigável e com potencial de mercado, adaptável às diferentes instituições do Sistema Nacional de Inovação, com base nas avaliações dos especialistas. Adicionalmente, a implementação da solução tecnológica permitiu observar características específicas da Rede Federal, como a diversidade de disciplinas e a concentração geográfica das pesquisas, bem como a necessidade de fortalecer a colaboração entre pesquisadores e o setor privado. A contribuição científica residiu na proposição de requalificação dos dados através do processo de extração, tratamento e carga, *data wherehouse*, classificadores como análise hierárquica e não hierárquica, identificação de fatores e indicadores para a inovação, permitindo inclusive de acréscimo no futuro de análises estatísticas mais complexas. Enquanto ilustração, após uma década, instituições do Nordeste e do Norte começaram a ganhar destaque. As conclusões reforçaram a necessidade de soluções tecnológicas práticas para os setores envolvidos e a importância da continuidade da pesquisa, sugerindo o aprofundamento teórico e a expansão metodológica, como a adoção de *data lakes* ou *lakehouse*. As limitações do estudo incluíram a dificuldade de acesso a dados públicos no Brasil e a necessidade de atualizações constantes do sistema para manter a qualidade e acessibilidade dos dados.

**Palavras-chave:** *Business Intelligence* (BI); *Dashboard e Data Warehouse*; *Triple Helix of Innovation* (TH); Canais de Pesquisa, Desenvolvimento Tecnológico e Serviços (PD&S); Instituições de Ensino Profissional, Científico e Tecnológico ou Superior

## ABSTRACT

Given the growing importance of addressing institutional, productive, and governmental demands in the Science, Technology, and Innovation ecosystem, this thesis sought to solve a critical problem: how to organize and enhance the management of research, technological development, and service provision (RD&S) data in Science and Technology Institutions (STIs) in order to respond in a timely and reliable manner to these demands. To this end, the work presented a developed online technological solution aimed at integrating and systematizing public data from this ecosystem. This was registered at the INPI (BR5120230010197), based on theories of innovation, business analysis and data visualization. It used the incorporation of Business Intelligence techniques, data warehouse and interactive dashboards as a quantitative approach. To validate this thesis, data from the Federal Network of Professional, Scientific, and Technological Education was used as a case study. The challenge was to overcome the barriers to integrating data related to innovation management, which has hindered decision-making by the general public, STI managers, researchers, and the productive sectors. By designing and implementing such technological solutions, anchored in the theories of Business Analysis, Data Visualization, and the Triple Helix Theory of Innovation, it was possible to present and better understand the problems from the availability of data to society in a timely and reliable manner. The methodology used was a quantitative and empirical-deductive approach that included literature review, consultation with users, identification data sources, and validation of the system by nine experts. The results showed that the tool was essential to consolidate information from multiple sources, helping society in general to respond quickly and reliably to the demands presented. The proposed solution also presented a user-friendly interface with market potential, adaptable to different institutions of the National Innovation System, based on expert assessments. In addition, the implementation of the technological solution allowed for the observation of specific characteristics of the Federal Network, such as the diversity of disciplines and the geographical concentration of research, as well as the need to strengthen collaboration between researchers and the private sector. The scientific contribution resided in the proposal of requalifying data through the process of extraction, treatment, and loading (ETL), data warehouse, classifiers such as hierarchical and non-hierarchical analysis, identification of factors and indicators of innovation, even allowing for the addition of more complex statistical analyses in the future. As an illustration, after a decade, institutions from the Northeast and North of Brazil began to gain prominence. The conclusions reinforced the need for practical technological solutions for the sectors involved and the importance of continued research, suggesting theoretical deepening and methodological expansion, such as the adoption of data lakes or lakehouse. Limitations of the study included the difficulty of accessing public data in Brazil and the need for constant system updates to maintain data quality and accessibility.

**Keywords:** *Business Intelligence (BI); Dashboard e Data Warehouse; Triple Helix of Innovation (TH); Research, Technology Development, and Services Channels (RTD&S); Federal Network of Professional, Scientific, and Technological Education (RFEPCT)*

## LISTA DE PUBLICAÇÕES

NASCIMENTO, H. H. S., SANTOS, C. N., ALVES, C. H. A., Nascimento, M. L. F. The X-patents. *World Patent Information*, v. 53, p. 1-13, 2018.

SGARBI, V. S., SANTOS, C. N., NASCIMENTO, M. L. F., LUIS, S. M. B., SGARBI, E. S., ALVES, C. H. A. Pós-graduações brasileiras em engenharia e a formação pedagógica docente: um estudo dos dados na Plataforma Sucupira. *Avaliação: Revista da Avaliação da Educação Superior (Campinas)*, v. 27, p. 91-117, 2022.

SGARBI, V. S., ALVES, C. H. A., SANTOS, C. N., NASCIMENTO, M. L. F., SOUZA, M. E. PG *Analytics*. Titulares IFAL e UFBA. Linguagem DAX. Campo de Aplicação AD-01, AD-04, FN-01, IF-01, IN-01. GI-01, TC-01. Registro de Software: BR 51 2023 002411 2, 2023.

NAUBER, C. N., SGARBI, V. S. NASCIMENTO, M. L. F., MORAES Jr., E. C., SOUZA, M. E. PD&S *Analytics*. Titulares IFAL e UFBA. Linguagem DAX. Campo de Aplicação AD-01, AD-04, FN-01, IF-01, IN-01. GI-01, TC-01. Registro de Software: BR 51 2023 001019 7, 2023.

## 1 Sumário

1	Sumário .....	13
2	Lista de Figuras .....	16
3	Lista de Tabelas.....	18
4	Lista de Quadros .....	19
5	Lista de Gráficos .....	21
6	Lista de Diagrama .....	22
7	Lista de Abreviaturas e Siglas.....	23
1	Introdução .....	27
1.1	Justificativa .....	33
1.2	Problema .....	37
1.3	Questão de Pesquisa .....	37
1.4	Tema... ..	38
1.5	Tese.....	38
1.6	Hipótese .....	38
1.7	Objetivo Geral .....	38
1.8	Objetivos Específicos .....	39
1.9	Contribuições Esperadas.....	39
2	Referencial Teórico.....	41
2.1	Revisão dos Resultados relacionados aos Canais de Pesquisa, Desenvolvimento Tecnológico e Prestação de Serviços da RFEPCT .....	46
2.2	Desafios e Soluções para a Interação ICTs-SPP para o Desenvolvimento Socioeconômico .....	54
2.3	Revisão da Evolução da Teoria Tripla Hélice de Inovação.....	58
2.4	Revisão da Literatura sobre <i>Business Intelligence, Data Warehouse,</i> <i>Dashboards</i> e Indicadores-chave de Desempenho Aplicados ao Ecossistema de ICTs 63	
2.4.1	Histórico e Evolução de BI.....	64
2.4.2	Teorias Relevantes em BI.....	66
2.4.3	Definição e Componentes de BI.....	68
2.4.4	Interações entre BI e a Revisão da Literatura.....	76
3	Metodologia .....	81
3.1	Seleção e Análise dos Estudos Utilizados neste Trabalho .....	83
3.2	Levantamento de Requisitos dos Usuários .....	86
3.3	Desenvolvimento do Modelo de Dados.....	88
3.4	Identificação de Fontes de Dados e Variáveis Identificados no Referencial Teórico Analisado .....	90
3.5	Modelo de BI .....	91
3.5.1	Fonte de Dados .....	93
3.5.2	CrITÉrios para Inclusão e Exclusão dos Dados.....	95
3.5.3	Extração, Transformação e Carga dos Dados Seleccionados para o Projeto .....	95
3.5.4	Modelo de Entidade-Relacionamento dos Dados.....	96

	3.5.5 .Implementação Híbrida de <i>Data Warehouse</i> : integrando Modelos Estrela e Floco de Neve para Otimização de Desempenho e Análise .....	96
	3.5.6 Tabelas Dimensão e Fato .....	97
	3.5.7 Indicador-chave de Desempenho (KPI).....	99
	3.5.8 <i>Dashboard</i> .....	100
	3.5.9 <i>Data Mining</i> .....	101
	3.6 Validação da Solução .....	105
4	Resultados .....	107
	4.1 Comparando Alguns Resultados da Educação Profissional e Tecnológica: um Estudo de Caso da RFEPCT Brasileira e Redes em Países como Alemanha, Canadá, Chile, China e Finlândia.....	108
	4.2 Resultado do Desenvolvimento do MVP .....	109
	4.2.1 Passo 1: Integração dos Dados .....	109
	4.2.2 Passo 2: Modelagem dos Dados .....	110
	4.2.3 Passo 3: Definição de Métricas e KPIs.....	110
	4.2.4 Passo 4: Criação do Painel Visual .....	112
	4.3 Resultados Gerados Pelo MVP Proposto .....	113
	4.4 Resultados obtidos sobre a Questão de Pesquisa e Hipótese Investigada .	123
	4.5 Resultados Descritivos da RFEPCT e da Revisão da Literatura .....	124
	4.5.1 Sinergia de Dados: Resultados da Análise Descritiva e Revisão da Literatura.....	126
	4.5.2 Investimentos e Distribuições.....	127
	4.5.3 Correlações .....	128
	4.5.4 Indicadores de Desempenho .....	129
	4.5.5 Análise Descritiva.....	129
	4.5.6 Análise de <i>Outliers</i> (Uni)Multivariada .....	131
	4.6 Resultados da Mineração dos Dados da RFEPCT.....	131
	4.6.1 Resultados das Análises HC e NHC .....	131
	4.6.2 Resultados da Análise da AFE por PCA .....	134
5	Análise e Discussão.....	139
	5.1 Análise Comparativa da Educação Profissional e Tecnológica no Brasil e em Outros Países: Semelhanças e Diferenças na Organização, Regulação e Financiamento.....	141
	5.2 Análise da Implementação de Tecnologias Específicas em Ecossistema das ICTs no Contexto da Teoria TH: Análise Crítica e Identificação de Desafios e Oportunidades.....	141
	5.3 Análise da Relevância e Desafios do Uso de <i>Dashboards</i> em ecossistema das ICTs.....	144
	5.4 Análise da Questão de Pesquisa e das Hipóteses Apresentadas .....	146
	5.5 Análise dos Resultado Descritivos da Rede Federal (2008-2018), da Revisão da Literatura e da Mineração de Dados.....	149
	5.6 Análise Conjuntural da Mineração de Dados Aplicada à Rede Federal (2008-2018).....	155

5.7	Análise dos Canais de Pesquisa, Desenvolvimento Tecnológico e Serviços em Ecossistema da Rede de Formação e Pesquisa em Ciência e Tecnologia: um Estudo de Caso Aplicado à Plataforma Proposta.....	167
5.7.1	Estudo de Caso Aplicado: Análise do Cenário Nacional de CT&I e PD&I na Primeira Década de Existência da RFEPCT .....	172
5.7.2	Estudo de Caso Aplicado: o Impacto das Características Individuais dos Pesquisadores nas Interações entre ICTs-SPP .....	184
6	Conclusões .....	188
7	Propostas para Futuros Trabalhos .....	190
8	Limitações Dessa Tese.....	192
9	Referências.....	193
10	APÊNDICES.....	220
	Apêndice A - Histórico da Rede Federal de Educação Profissional, Científica e Tecnológica .....	220
	Apêndice B - Quadros, Tabelas etc.....	226
	Apêndice C - Métodos de Interdependência Multivariada Exploratória ou Mineração de Dados .....	302
	C.1 - Análise <i>Cluster</i> Hierárquico (HC) e Não Hierárquico (NHC).....	302
	C.2 – Análise Fatorial Exploratória (EFA) por Componentes Principais (PCA) .....	314
	Apêndice D - Detalhamento Estrutural das Tabelas de Dimensão e Fato empregadas na Construção do <i>Data Warehouse</i> .....	322
	Apêndice E - Detalhamento do Processo de ETL .....	324
	E.1 - Raspagem dos Dados. ....	324
	E.2 - Limpeza .....	326
	E.3 - Duplicação.....	326
	E.4 – Revisão de Formato .....	327
	E.5 – Reestruturação de Chaves.....	327
	E.6 – Filtragem.....	327
	E.7 – Unindo Dados .....	328
	E.8 – Dividindo Dados.....	328
	E.9 – Validando Dados .....	329
	E.10 – Sumarização.....	329
	E.11 – Agregação .....	330
	E.12 – Integração .....	330
	Apêndice F - Detalhamento da Concepção do Projeto de Design UX para os Três Módulos ( <i>Dashboard</i> ) propostos para o MVP .....	331
	Apêndice G - Detalhamento do Processo de Validação do Projeto de <i>UX Design</i> para os Três Módulos ( <i>Dashboard</i> ) propostos para o MVP .....	336
	Apêndice H - Simulação de Acesso ao MVP: Resultados do Usuário W de uma ICT.....	341
	Apêndice I - Modelo de Documento encaminhado para os Avaliadores .....	347

## 2 Lista de Figuras

- Figura 1: Representação esquemática da técnica clássica de BI a partir de fontes de dados. Adaptado das referências Krmac (2011) e Fana *et al.* (2021), os *cheks* positivos identificam os resultados de BI adotados neste trabalho..... 91
- Figura 2: Representação esquemática do modelo de BI utilizado neste trabalho, a partir de dados públicos, adaptada de Krmac (2011) e Fana *et al.* (2021). ..... 92
- Figura 3: Ilustração das tabelas dimensões (representada pelas áreas de linha tracejada, filtros) e fato (representada pela área de linha traço dois pontos, métricas) para o período de 2008 a 2018 da RFEPCT, disponibilizado no MVP. .... 98
- Figura 4: Visão geral das funcionalidades do dashboard MVP em diferentes módulos. .... 116
- Figura A. 1: Mapa territorial da Rede Federal de Educação Profissional, Científica e Tecnológica demonstrando sua abrangência e capilaridade até 2022..... 223
- Figura B. 1: Representação visual do modelo de entidade de relacionamento (MER) utilizado para criar o Data Warehouse (DW) do projeto. .... 289
- Figura B. 2: Esquema de uma pequena fração do banco de dados temporário do modelo de entidade de relacionamento utilizada para a carga do Data Warehouse ..... 290
- Figura B. 3: Apresentação do procedimento de ETL aplicado neste trabalho, adaptado de adaptado de Paikaray et al. (2021, p. 505-508)..... 290
- Figura B. 4: Esquema Geral para Carregar os dados no Power BI ..... 291
- Figura B. 5: Seleção do tipo de banco de dado padrão MySQL ..... 291
- Figura B. 6: Acesso ao banco de dados do MySQL para carregar os dados no Power BI ..... 291
- Figura B. 7: Tabelas de dados do MySQL carregadas no Power BI no início do projeto. Região tracejada destaca tabelas pré-tratadas no MySQL, enquanto região com dois pontos indica falta de modelagem dimensional no Data Warehouse do Power BI ..... 292
- Figura B. 8: Ilustração das tabelas de dados pós-tratamento no Power BI (segundo ETL). Região tracejada mostra tabelas pós-tratamento. Região com dois pontos destaca várias modelagens dimensionais (tabelas de fatos e dimensões) implementadas no DW no final do projeto..... 292
- Figura B. 9: Ilustração da implementação da modelagem dimensional do esquema floco de neve da métrica Patente ( $X_{20}$  e  $X_{23}$ ) no Data Warehouse no Power BI. A região tracejada mostra as tabelas fato e dimensões organizadas. A região com dois pontos destaca a aba correspondente à modelagem dimensional no projeto ..... 293
- Figura B. 10: Ilustração da implementação da codificação de uma métrica no *Data Warehouse* do Power BI com DAX ..... 293
- Figura B. 11: Ilustração da implementação da métrica "Total de Orientação de Pesquisa" utilizando DAX no *Data Warehouse* do Power BI. A região tracejada representa a visualização final e a com dois pontos indica a extraída da modelagem de dados..... 294
- Figura B. 12: Ilustração da implementação da codificação de um KPIs no Data Warehouse do Power BI com DAX ..... 294



Figura B. 13: Ilustra a página inicial que contém a seleção das duas Tecnologias Desenvolvidas pelo Grupo de Pesquisa NUMDI: esquerda registro BR 51 2023 001019 7 (INPI, o Autor, 2023) e a direita registro BR 51 2023 002411 2 (INPI, Sgarbi, 2023). .....	295
Figura B. 14: Ilustração do UX design do MVP proposto neste estudo registrado sob o registro BR 51 2023 001019 7 (INPI).....	295
Figura B. 15: Ilustração do UX <i>design dashboard</i> para o módulo – <b>Amplo</b> .....	296
Figura B. 16: Ilustração do UX design dashboard para o módulo – <b>Detalhado</b> .....	297
Figura B. 17: Ilustração do UX <i>design dashboard</i> para o módulo – <b>Comparativo</b> ....	298
Figura B. 18: Estudo comparativo de Agrupamento Hierárquico das Abordagens 1e 2 para o intervalo temporal entre 2008 e 2018 da RFEPCT. ....	300
Figura B. 19: Estudo comparativo de Agrupamento Não Hierárquico (NHC: <i>K-means</i> ) das Abordagens 1e 2 para o intervalo temporal entre 2008 e 2018 da RFEPCT..	301
Figura C. 1: ilustra a distribuição inicial das observações (Int.0). A etapa Int.1 mostra a seleção arbitrária dos centroides iniciais. As etapas Int.1 a Int.4 demonstram as iterações do método <i>K-means</i> , utilizando o algoritmo de otimização Hartigan e Wong, com o objetivo de identificar os agrupamentos ideais. A etapa Int.5 destaca o resultado do agrupamento ideal. Computacionalmente, este resultado ilustra o agrupamento NHC, específico para o Cenário 12, na Abordagem 2. ....	313
Figura E. 1:Representação do esquema floco de neve apresentando o relacionamento entre as tabelas dimensão e fato para a elaboração do dashboard - Módulo atividade PD&S: <b>Amplo</b> .....	323
Figura E. 2:Representação do esquema floco de neve apresentando o relacionamento entre as tabelas dimensão e fato para a elaboração do dashboard - Módulo atividade PD&S: <b>Detalhado</b> .....	323
Figura E. 3:Representação do esquema floco de neve apresentando o relacionamento entre as tabelas dimensão e fato para a elaboração do dashboard - Módulo atividade PD&S: <b>Comparativo</b> .....	323
Figura F. 1:Esquema Geral de <i>UX design</i> do dashboard MVP .....	333

### 3 Lista de Tabelas

Tabela B. 1: Somatório dos dados referentes aos canais de PD&S das 42 unidades da Rede Federal Brasileira de Educação Profissional, Científica e Tecnológica entre 2008 e 2018. A descrição dos dados está disponível no Apêndice B, Quadro B. 14: Região ( $W_1$ ), Região ( $UF_2$ ), Órgão ( $W_3$ ), Investimento público em Pesquisa e Desenvolvimento ( $X_1$ ), Pesquisadores ( $X_2$ ), Orientações de Pesquisa ( $X_3$ ), Projetos de Pesquisa ( $X_4$ ), Linhas de Pesquisa ( $X_5$ ), Publicações científicas ( $X_6$ ), Pesquisas financiadas pelos setores produtivos ( $X_7$ ), Linhas de pesquisa com setores produtivos ( $X_8$ ), Projetos de Pesquisas financiadas pelos setores produtivos ( $X_9$ ), Treinamento ( $X_{10}$ ), Consultoria ( $X_{11}$ ), Projetos de extensão acadêmicos ou tecnológicos ( $X_{12}$ ), Serviços de uso de laboratórios e equipamentos; etc. ( $X_{13}$ ), Treinamentos demandados pelos setores produtivos ( $X_{14}$ ), Consultorias demandadas pelos setores produtivos ( $X_{15}$ ), Projetos de extensão demandados pelos setores produtivos ( $X_{16}$ ), Serviços de uso de laboratórios e equipamentos; etc. demandados pelos setores produtivos ( $X_{17}$ ), Marcas (concedidas) ( $X_{18}$ ), Desenhos Industriais (concedidos) ( $X_{19}$ ), Patentes (publicadas) ( $X_{20}$ ), Programas de computador (concedidos) ( $X_{21}$ ), Desenhos Industriais (concedidos) demandados pelos setores produtivos ( $X_{22}$ ), Patentes (publicadas) demandadas pelos setores produtivos ( $X_{23}$ ) e Programas de computador (concedidos) demandados pelos setores produtivos ( $X_{24}$ ). ..... 256	
Tabela B. 2: Estatísticas descritivas dos dados referentes aos canais de PD&S das 42 unidades da Rede Federal Brasileira de Educação Profissional, Científica e Tecnológica entre 2008 e 2018 ..... 262	
Tabela B. 3: Análise das correções dos dados referentes as 24 variáveis relacionada aos canais de PD&S das 41 unidades da Rede Federal Análise das correções dos dados referentes as 24 variáveis relacionada aos canais de PD&S das 41 unidades da Rede Federal Brasileira de Educação Profissional, Científica e Tecnológica (outlier: UTFPR), entre 2008 e 2018 ..... 263	
Tabela B. 4: Análise dos Principais Componentes das Variáveis $X_1$ a $X_{24}$ ..... 266	
Tabela B. 5: Resultado da Data Mining Integrada de Canais de PD&S em ICTs 2008 a 2012 (Cenário 12, Abordagem 1)..... 270	
Tabela B. 6: Análise Comparativa da Evolução e Posicionamento do Brasil em Indicadores de Pesquisa e Inovação: 2008-2018..... 288	
Tabela C. 1: Dados Simulados para $A_1$ (em unidades arbitrárias) ..... 318	
Tabela C. 2: Dados Simulados para $A_2$ (em unidades arbitrárias) ..... 318	
Tabela D. 1: Exemplo da tabela patentes (INPI), extraída utilizando a linguagem $R$ , contendo os dados brutos no formato linha x coluna utilizado para inserção no banco de dados temporário ao final do processo de ETL ..... 326	

## 4 Lista de Quadros

Quadro B. 1: Apresentação das principais plataformas públicas e privadas disponíveis até 2023 que buscam organizar os dados administrativos, financeiros, educacionais, de pesquisa e de extensão relacionados disponibilizados pelas ICTs.....	226
Quadro B. 2: Comparação das Principais Teorias que foram propostas para investigar os diversos tipos de interação entre as ICTs-SPP, com o objetivo de avaliar a natureza e o impacto dessas relações.....	227
Quadro B. 3: Avaliação original dos especialistas em relação ao MVP proposto.....	229
Quadro B. 4: Síntese Evolutiva do Modelo Tripla Hélice 1958 – 2022.....	238
Quadro B. 5: Principais variáveis utilizadas para analisar os Canais de PD&S em ICTs no contexto da Teoria TH.....	240
Quadro B. 6: Principais diferenças entre esquemas estrela e floco de neve.....	243
Quadro B. 7: Resultado da revisão bibliométrica utilizado para elaboração do histórico da RFEPCT .....	244
Quadro B. 8: Resultado da revisão bibliométrica utilizado para elaboração do histórico da evolução dos canais de PD&S do ecossistema da RFEPCT .....	244
Quadro B. 9: Resultado da revisão bibliométrica utilizado para contextualizar alguns fatores nacionais de CT&I e PD&I em que o ecossistema das ICTs se desenvolveu nas duas últimas décadas.....	245
Quadro B. 10: Resultado da revisão bibliométrica utilizado para contextualizar a teoria da TH da inovação utilizada na investigação dos fenômenos dos canais de PD&S desenvolvidos no ecossistema das ICTs .....	245
Quadro B. 11: Resultado da revisão bibliométrica utilizado para contextualizar a teoria da tripla hélice da inovação utilizada na investigação dos fenômenos dos canais de PD&S desenvolvidos no ecossistema das ICTs .....	246
Quadro B. 12: Síntese das fontes de dados públicas brasileiras identificados geralmente em estudos canais de PD&S, a partir do referencial teórico (Seção 2).....	247
Quadro B. 13: Expansão das fontes de dados oficiais brasileiras consideradas (Quadro B. 12), incluindo rótulo e sigla.....	247
Quadro B. 14: Características das 28 variáveis/métricas consideradas para a coleta dos dados referentes aos canais de PD&S. ....	248
Quadro B. 15: Resumo das Variáveis Utilizadas e Suas Categorizações no Estudo....	255
Quadro B. 16: Lista dos códigos ou atividades econômicas na CNAE.....	280
Quadro B. 17: Comparação de Funcionalidades e Disponibilidade de Informações: Análise do Dashboard Proposto versus Portal Integra, Portal Universidade 360 e StelaExperta® .....	281
Quadro B. 18: Estudos das variáveis contributivas para a separação dos agrupamentos HC, Cenário 12, Abordagem 1 e Abordagem 2 utilizando os testes de ANOVA, Benjamini_Hochberg (BH), Bonferroni e Magnitude.....	283

Quadro B. 19: Síntese da Análise Hierárquica de Agrupamentos (HC) das ICTs por Indicadores para o Cenário 12, Abordagem 1 .....	284
Quadro B. 20: Síntese da Análise Hierárquica de Agrupamentos (HC) das ICTs por Indicadores para o Cenário 12, Abordagem 2.....	285
Quadro B. 21: Síntese da Análise Não Hierárquica de Agrupamentos (NHC: <i>K-means</i> ) das ICTs por Indicadores para o Cenário 12, Abordagem 1 .....	286
Quadro B. 22 Síntese da Análise Não Hierárquica de Agrupamentos (NHC: <i>K-means</i> ) das ICTs por Indicadores para o Cenário 12, Abordagem 2 .....	287
Quadro C. 1: Comparação entre Algoritmos de Otimização para <i>K-means</i> .....	311

## **5 Lista de Gráficos**

Gráfico B. 1: Evolução dos investimentos em P&D entre 2008 e 2018 na RFEPCT.. 299

## **6 Lista de Diagrama**

Diagrama 1: Esquema metodológico aplicado nesta tese.....	82
---	----

## 7 Lista de Abreviaturas e Siglas

AN	- Teorias da Análise de Negócios
APA	- <i>American Psychological Association</i> (Associação Americana de Psicologia)
API	- Interface de Programação de Aplicativos
ARI	- Índice Ajustado <i>Rand</i>
AVA	- Ambiente Virtual de aprendizagem
BD	- Bancos de Dados
BES	- <i>Business Enterprise Sector</i>
BI	- <i>Business Intelligence</i>
BSC	- <i>Balanced Scorecard</i>
C&T	- Ciência e Tecnologia
CA	- <i>Cluster Analysis</i> (Análise de Agrupamentos)
CAPES	- Coordenação de Aperfeiçoamento de Pessoal de Nível Superior
CEFET	- Centro Federal de Educação Tecnológica
CEFET/MG	- Centro Federal de Educação Tecnológica de Minas Gerais
CEFET/RJ	- Centro Federal de Educação Tecnológica do Rio de Janeiro
CES	- Censo da Educação Superior
CGEE	- Centro de Gestão e Estudos Estratégicos, Organização Social
CGU	- Tribunal de Contas da União
CNAE	- Classificação Nacional e Atividades Econômicas do Brasil
CNPJ	- Cadastro Nacional de Pessoa Jurídica
CNPq	- Conselho Nacional de Desenvolvimento Científico e Tecnológico
CPII	- Colégio Pedro Segundo
CT&I	- Ciência, Tecnologia e Inovação
CV	- Coeficiente de Variação
DASHBOARD	- Painel Gráfico
DAX	- <i>Data Analysis Expressions</i>
DE	- Dedicção Exclusiva para o serviço público
DGP	- Diretório do Grupo de Pesquisa
DM	- <i>Data Mining</i> (Mineração de Dados)
DT	- Desenvolvimento Tecnológico
DW	- <i>Data Warehouse</i> (Banco de Dados Multidimensional Otimizado)
EAD	- Educação a Distância
EBTT	- Ensino Básico, Técnico e Tecnológico
EFA	- <i>Exploratory Factor Analysis</i> (Análise Fatorial Exploratória)
EJA	- Ensino de Jovens e Adultos
EMBRAPA	- Empresa Brasileira de Pesquisa Agropecuária
EPCT	- Educação Profissional, Científica e Tecnológica
ETL	- Extração, Carga e Tratamento
FAP	- Fundações de Apoio e Pesquisa
FAPs	- Fundações de Amparo à Pesquisa
FIC	- Formação Inicial Continuada
Finep	- Financiadora de Estudos e Projetos
FORPOG	- Fóruns dos Pró-reitores de Pesquisa e Inovação
GC	- Gestão do Conhecimento
GERD	- <i>Gross Expenditure on R&amp;D</i> (Despesa Bruta em Pesquisa e Desenvolvimento)
GI	- Gestão da Informação

GII	- <i>Global Innovation Index</i> (Índice de Inovação Global)
GOV	- Governo
HC	- <i>Hierarchical Clustering</i> (Agrupamento Hierárquico)
IBGE	- Instituto Brasileiro de Geografia e Estatística
IBC	- Índice Baseado em Componentes Principais
ICTs	- Instituições Científicas, Tecnológicas e de Inovação
ICT-MACE	- Melhores ICTs Avaliadas pelo Critério Estabelecido
IDHM	- Índice de Desenvolvimento Humano Municipal
IEPCT	- Instituições de Ensino, Pesquisa, Ciência e Tecnologia
IF	- Instituto Federal
IFAC	- Instituto Federal de Educação, Ciência e Tecnologia do Acre
IFAL	- Instituto Federal de Educação, Ciência e Tecnologia de Alagoas
IFAM	- Instituto Federal de Educação, Ciência e Tecnologia do Amazonas
IFAP	- Instituto Federal de Educação, Ciência e Tecnologia do Amapá
IFB	- Instituto Federal de Educação, Ciência e Tecnologia de Brasília
IFBA	- Instituto Federal de Educação, Ciência e Tecnologia da Bahia
IFBaiano	- Instituto Federal de Educação, Ciência e Tecnologia Baiano
IFC	- Instituto Federal de Educação, Ciência e Tecnologia Catarinense
IFCE	- Instituto Federal de Educação, Ciência e Tecnologia do Ceará
IFES	- Instituto Federal de Educação, Ciência e Tecnologia do Espírito Santo
IFF	- Instituto Federal de Educação, Ciência e Tecnologia Fluminense
IFFAR	- Instituto Federal de Educação, Ciência e Tecnologia Farroupilha
IFGO	- Instituto Federal de Educação, Ciência e Tecnologia de Goiás
IFGoiano	- Instituto Federal de Educação, Ciência e Tecnologia Goiano
IFMA	- Instituto Federal de Educação, Ciência e Tecnologia do Maranhão
IFMG	- Instituto Federal de Educação, Ciência e Tecnologia de Minas Gerais
IFMS	- Instituto Federal de Educação, Ciência e Tecnologia de Mato Grosso do Sul
IFMT	- Instituto Federal de Educação, Ciência e Tecnologia de Mato Grosso
IFNMG	- Instituto Federal de Educação, Ciência e Tecnologia do Norte de Minas Gerais
IFPA	- Instituto Federal de Educação, Ciência e Tecnologia do Pará
IFPB	- Instituto Federal de Educação, Ciência e Tecnologia da Paraíba
IFPE	- Instituto Federal de Educação, Ciência e Tecnologia de Pernambuco
IFPI	- Instituto Federal de Educação, Ciência e Tecnologia do Sertão do Piauí
IFPR	- Instituto Federal de Educação, Ciência e Tecnologia do Paraná
IFRJ	- Instituto Federal de Educação, Ciência e Tecnologia do Rio de Janeiro
IFRN	- Instituto Federal de Educação, Ciência e Tecnologia do Sertão do Rio Grande do Norte
IFRO	- Instituto Federal de Educação, Ciência e Tecnologia de Rondônia
IFRR	- Instituto Federal de Educação, Ciência e Tecnologia de Roraima
IFRS	- Instituto Federal de Educação, Ciência e Tecnologia do Rio Grande do Sul



IFSC	- Instituto Federal de Educação, Ciência e Tecnologia de Santa Catarina
IFSE	- Instituto Federal de Educação, Ciência e Tecnologia de Sergipe
IFSEMG	- Instituto Federal de Educação, Ciência e Tecnologia do Sudeste de Minas Gerais
IFSMG	- Instituto Federal de Educação, Ciência e Tecnologia do Sul de Minas Gerais
IFSP	- Instituto Federal de Educação, Ciência e Tecnologia de São Paulo
IFSTPE	- Instituto Federal de Educação, Ciência e Tecnologia do Sertão do Pernambuco
IFSUL	- Instituto Federal de Educação, Ciência e Tecnologia Sul-Rio-Grandense
IFTM	- Instituto Federal de Educação, Ciência e Tecnologia do Triângulo Mineiro
IFTO	- Instituto Federal de Educação, Ciência e Tecnologia do Tocantins
IGCs	- Instituições de Gestão do Conhecimento
IIP_PD&S	- Índice Integrado de Performance e Inovação em Pesquisa, Desenvolvimento Tecnológico e Prestação de Serviços Técnicos Especializados
INEP	- Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira
INPI	- Instituto Nacional da Propriedade Industrial
IPT	- Instituto de Pesquisas Tecnológicas
KPI	- Indicador-chave de Desempenho
MCTI	- Ministério da Ciência, Tecnologia e Inovações
MEC	- Ministério da Educação e Cultura
MER	- Modelo de Entidade de Relacionamento
MPF	- Ministério Público Federal
MSAI	- Medida de Adequação da Amostra Individual
MVP	- <i>Minimum Viable Product</i> (Mínimo Produto Viável)
NACE	- Nomenclatura de Atividades Econômicas da União Europeia
NAICS	- Sistema de Classificação da Indústria da América do Norte
NHC	- <i>Non Hierarchical Clustering</i> (Agrupamento Não Hierárquico)
NIT	- Núcleo de Inovação Tecnológica
NUMDI	- Núcleo Multidisciplinar de Inovação Tecnológica
OCDE	- Organização para a Cooperação e Desenvolvimento Econômico
OLAP	- <i>Online Analytical Processing</i> (Processamento Analítico <i>Online</i> )
P&D	- Pesquisa e Desenvolvimento Tecnológico
PACTI	- Plano de Ação 2007–2010 para Ciência, Tecnologia e Inovação
PCA	- <i>Principal Component Analysis</i> (Análise de Componentes Principais)
PD&I	- Pesquisa, Desenvolvimento Tecnológico e Inovação
PD&S	- Pesquisa, Desenvolvimento Tecnológico e Serviços
PIB	- Produto Interno Bruto
PISA	- Programa Internacional de Avaliação de Estudantes
PROEJA	- Programa Nacional de Integração da Educação Profissional com a Educação Básica na Modalidade de Educação de Jovens e Adultos
PVS	- <i>Public Value Scorecard</i>
Q&A	- Filtro Inteligente do <i>Power BI</i>
R&D	- <i>Research and Development</i> (Pesquisa e Desenvolvimento)

REM	- <i>Relational Entity Model</i>
RFEPCT	- Rede Federal de Educação Profissional, Científica e Tecnológica
SESU	- Secretaria de Educação Superior
SETC	- Secretaria de Educação Profissional e Tecnológica
SGDB	- Sistema Gerenciador de Banco de Dados
SIC-CD	- Serviço de Informação ao Cidadão
SIG	- Sistema de Informação Gerencial
SNCTI	- Sistema Nacional de Ciência, Tecnologia e Inovação
SNI	- Sistema Nacional de Inovação
SPP	- Setores Públicos e Privados
SQL	- <i>Structured Query Language</i> (Linguagem de Consulta Estruturada)
TAN	- Teoria da Análise de Negócios
TCU	- Tribunal de Contas da União
TD	- Tabela Dimensão
TDR	- Teoria da Decisão Racional
TF	- Tabela Fato
TGC	- Teoria da Gestão do Conhecimento
TH	- Tripla Hélice de Inovação
TI	- Tecnologia da Informação
TIC	- Teoria da Inteligência Competitiva
TT	- Transferência Tecnológica ou de Tecnologia
TTO	- <i>Technology Transfer Office</i> (Escritório de Transferência Tecnológica)
TVD	- Teoria da Visualização de Dados
UF	- Unidade da Federação
UIC	- Colaboração Universidade-Indústria
UIG	- Universidade-Industria-Governo
UTFPR	- Universidade Federal Tecnológica do Paraná
VD	- Visualização de Dados
YOY	- <i>Year Over Year</i> (Ano a Ano)

## 1 Introdução

Atualmente a gestão de dados públicos, sob a política de dados abertos dos canais de pesquisa, desenvolvimento tecnológico e serviços (PD&S) das Instituições Científicas, Tecnológicas e de Inovação (ICTs) nunca foi tão crítico, pois na atualidade conhecimento é poder. Ela é vista como um pilar essencial para promover a transparência e fomentar a inovação.

Definida pela Open Knowledge Brasil<sup>1</sup> com dados que podem ser utilizados, realizados e redistribuídos pela sociedade, desde que as exigências legais e das licenças que os dados contidos nas diversas fontes originais sejam compartilhados. Este quadro é particularmente relevante no contexto do Sistema Nacional de Inovação, que é composto pela interação entre ICTs, Governo e Setores Produtivos, cujos resultados dependem crucialmente da capacidade de gerar conhecimento, desenvolver novas tecnologias, atuar no desenvolvimento socioeconômico, bem como em última instância atender às demandas dos diversos setores da sociedade em geral (gestores, pessoas jurídicas, pesquisadores, governo e pessoas físicas).

Nesse ecossistema, a política de dados abertos serve como um catalisador para a inovação colaborativa, possibilitando que o conhecimento gerado pelas ICTs possa ser amplamente acessado e utilizado para impulsionar o desenvolvimento tecnológico e socioeconômico, além de fortalecer as bases para respostas ágeis a desafios emergentes globais.

No entanto, a gestão dessa grande quantidade de dados gerados pelas ICTs permanece um desafio. No cenário brasileiro, esses desafios já vêm sendo mapeados por pesquisadores, pelos setores econômicos e por órgãos como o Tribunal de Contas da União (TCU) e o Ministério da Ciência, Tecnologia e Inovações (MCTI), dentre outros, a algum tempo. Esse mapeamento evidencia a necessidade premente de soluções inovadoras que possam superar as barreiras existentes no acesso, na integração, na análise de dados e na publicidade transparente dos resultados fundamentados em evidências concretas. Assim, será possível estimular a inovação e promover o desenvolvimento sustentável do país.

---

<sup>1</sup> “A *Open Knowledge* Brasil (OKBR), também chamada de Rede pelo Conhecimento Livre, é o capítulo da *Open Knowledge* Internacional no Brasil. Ela é uma organização da sociedade civil (OSC) sem fins lucrativos e apartidária, regida por um estatuto, decidido em assembleia realizada no dia 4 de setembro de 2013, sendo reconhecida em cartório no dia 3 de outubro de 2013”. Disponível em <https://ok.org.br/>, acesso em jul. 2014.

No Brasil, a complexa gestão dos dados das ICTs referentes aos canais de exige soluções tecnológicas avançadas. Reconhecido por pesquisadores e órgãos como o TCU e o MCTI, este desafio impulsiona o desenvolvimento de uma plataforma online baseada em Business Intelligence (BI) para unificar e clarificar dados de pesquisa e desenvolvimento. Tal solução busca melhorar a transparência e eficiência da gestão de dados, promovendo a inovação e o desenvolvimento sustentável.

Com o objetivo de fundamentar as decisões, análises, gestão e monitoramento por meio de evidências, utilizou-se das Teorias da Análise de Negócios (AN) que busca identificar e analisar necessidades e requisitos empresariais para aprimorar a eficiência de processos (Chasin, 2016), Visualização de Dados (VD) que busca transformar dados brutos em representações gráficas claras, facilitando a identificação de padrões e insights (Palloni et al., 2018), e Tripla Hélice de Inovação (TH) que busca explicar a dinâmica dos sistemas de inovação através da interação entre universidades, indústria e governo (Etzkowitz et al., 2000). Estas teorias serviram como referência para o levantamento de requisitos, desenvolvimento e proposição da solução tecnológica inédita, para o contexto considerado.

A investigação dos canais de PD&S objetivou identificar dados capazes de fornecer respostas rápidas aos gestores das ICTs e a sociedade em geral. Os gestores representam o grupo de líderes e administradores, responsáveis pela tomada de decisão, nas ICTs, Governo e Setores Produtivos. Dentre outras atividades, eles captam, alocam e administram os recursos, gerenciam projetos, asseguram a conformidade com regulamentações legais e promovem o compartilhamento dos resultados e captação das demandas relacionadas aos canais de PD&S.

A sociedade em geral precisa de soluções tecnológicas avançadas para superar os desafios relacionados ao acesso, integração e análise de dados, bem como de ferramentas que permitam a publicidade transparente dos resultados baseados em dados. Requerem, também, capacidade de analisar de forma simples a complexidade inerente aos ecossistemas do Sistema Nacional de Inovação. Para tal, faz-se necessário adotar técnicas de Business Intelligence (BI), Data Warehouses (DW), e outras tecnologias de processamento e análise de dados para alcançar tal finalidade.

Para superar os desafios e alcançar a finalidade apresentada, o desenvolvimento de soluções tecnológicas implementando as técnicas e ferramentas supracitadas são fundamentais para simplificar a análise da complexidade dos canais de PD&S, facilitando a tomada de decisões, o cumprimento de exigências legais e regulatórias, e a promoção

da eficiência em gestão de projetos e de seus recursos. Este enfoque é corroborado por estudos e relatórios de órgãos como o TCU e o MCTI, que sublinham os problemas na dispersão e integração de dados como barreiras ao progresso.

A criação de soluções tecnológicas em lidar com tais desafios não só melhoraria a gestão de informações nas ICTs, como também potencializar o desenvolvimento socioeconômico e a inovação no país, por meio das interações entre os atores da Tripla Hélice da inovação.

Diante desse contexto, em um primeiro momento, tal análise visou facilitar o atendimento às solicitações dos órgãos governamentais de controle. Posteriormente, buscou-se atender às demandas provenientes do governo, pesquisadores, sociedade e setores produtivos, tanto públicos quanto privados, monitorando e discernindo necessidades associadas à evolução desses canais ao longo do tempo.

As ICT são órgãos ou entidades da administração pública ou organizações sem fins lucrativos, legalmente constituídas sob as leis brasileiras, cuja missão, objetivo social ou estatutário inclui a execução de pesquisa básica ou aplicada, bem como o desenvolvimento de novos produtos, serviços ou processos (Brasil, 2016a - Art. 2º, Item V).

Neste contexto, também se encaixam as Instituições de Ensino Superior e a Rede Federal de Educação Profissional, Científica e Tecnológica (RFEPCT), bem como universidades e outras instituições de ensino com objetivos similares. Apesar dessas instituições de ensino possuírem seus objetivos específicos, elas compartilham com as ICT a promoção da inovação, do desenvolvimento econômico e do progresso do conhecimento. Portanto, neste estudo, o termo Instituições Públicas ou Sociais dedicadas ao Ensino, à pesquisa básica ou aplicada e ao desenvolvimento de novos produtos, serviços ou processos, doravante será representada pela sigla ICT.

Para investigar a aplicabilidade dessa abordagem, que busca integrar técnicas de BI, DW, e outras tecnologias de processamento e análise de dados dentro do contexto do Sistema Nacional de Inovação, relacionado aos canais de PD&S administrados pelas ICTs, optou-se por considerar os dados de PD&S armazenados e publicamente disponíveis relacionados ao ecossistema da RFEPCT. OS dados foram coletados entre 2008 e 2018.

Considerou-se os dados da Educação Profissional Brasileira, em especial àqueles referentes a RFEPCT considerando sua compreensão e contribuição histórica, que evoluiu e culminou na sua criação em 31 de dezembro de 2008 (Apêndice A).

Tal escolha deve-se aos objetivos e propósitos da RFEPCCT em fomentar o desenvolvimento regional, atendendo às demandas dos setores produtivos públicos e privados (SPP) e do governo (GOV). O período selecionado abarca a primeira década de funcionamento efetivo dessa rede de formação técnica e profissional, e a fundamentação teórica foi escolhida por ser interdisciplinar e possibilitar um estudo longitudinal.

Diversos órgãos, como o Tribunal de Contas da União (TCU) e o Ministério da Ciência, Tecnologia e Inovação (MCTI) ao elaborarem relatórios ou desenvolverem observatórios de dados, estes são evidências irrefutáveis de que o governo vem buscando identificar os problemas e apresentar algumas soluções tecnológicas, que possibilitem obter, conhecer, integrar, filtrar, gerenciar, disponibilizar e monitorar os canais de PD&S, entre outros. Essas demandas são evidentes, por exemplo:

- Nos relatórios do TCU (Brasil, 2019a) ao apresentar uma lista de riscos sobre a efetividade das políticas públicas de inovação
- No MCTI (Brasil, 2020), por exemplo ao desenvolver sistemas para monitorar a produção científica e tecnológica, bem como as tendências de inovação, fornecendo indicadores para orientar políticas públicas de CT&I no Brasil e sua avaliação em âmbito internacional através do Observatório de Ciência, Tecnologia e Inovação (OCTI), porém esse observatório não oferece uma interação multidimensional além de outras características como análises de correlação e cruzamento e filtro de dados multivariados em tempo real como a solução proposta neste trabalho.
- No setor privado, identificou-se apenas uma tecnologia *online* denominada Stela Experta, registrada sob o registro BR 51 2016 001253 6. Até o momento, não foram encontradas soluções privadas alternativas ou concorrentes, similares à proposta apresentada pelo SPP e as públicas são restritas aos canais de ensino, administração e finanças e em contexto muito específico e limitado abordando os Canais de Pesquisa, Desenvolvimento Tecnológico e Inovação (PD&I) e Ciência, Tecnologia e Inovação (CT&I) (Apêndice B, Quadro B. 1).

O governo brasileiro já tem aplicado a gestão pública de dados em plataformas como Nilo Peçanha (Brasil, 2018b), que busca gerenciar os canais de ensino, administração e finanças da RFEPCCT e Painel Universidade 360 (Brasil, 2021<sup>a</sup>), que busca gerenciar os dados abertos das Universidades Federais relacionados aos canais de CT& e PD&I, eixando de fora, por exemplo os canais de extensão e prestação de serviço.

A revisão de literatura permitiu identificar que a análise acadêmica e a visualização de dados no ensino superior são tópicos relevantes e em crescimento nas pesquisas (Stewart *et al.*, 2022), e a aplicação de *dashboards* são aspectos importantes a serem considerados (Wang *et al.*, 2015).

Este trabalho explora a expansão do escopo dos canais de ciência, pesquisa e inovação, para o de ciência, pesquisa, inovação, desenvolvimento tecnológico, extensão e prestação de serviços (PD&S) e suas inter-relações, e para a gestão integrada desses canais disponibiliza uma ferramenta tecnológica online pública mais completa e versátil para poder contribuir com as constantes demandas dos SPP e GOV, bem como da sociedade em geral.

Uma das principais limitações em integrar os canais de PD&S é a dificuldade no acesso a dados. Isso ocorre porque os dados de PD&S estão frequentemente fragmentados e não padronizados, conforme identificado pelo TCU no seu relatório de alto risco relacionado a Inovação e a gestão dos dados Públicos TCU (Brasil, 2019a). Como resultado, é difícil reunir informações sobre tais canais e avaliar efetivamente e objetivamente seu impacto. Outra limitação da abordagem atual é a limitação colaborativa entre os diversos atores envolvidos no desenvolvimento do ecossistema nacional de PD&I e de CT&I. Isso dificulta a identificação e a implementação de soluções inovadoras que atendam às demandas dos SPP, do GOV e das ICTs.

Este trabalho propõe uma abordagem para superar tais limitações. A solução tecnológica proposta adotou a técnica de BI associada às tecnologias de *Data Warehouse* (em inglês, DW) e *dashboard*, fundamentada nas teorias da Análise de Negócios, de Visualização de Dados e da Tripla Hélice da Inovação.

A adoção da técnica de BI pode ser uma ferramenta valiosa para as ICTs que buscam melhorar a tomada de decisão e a gestão. Ela pode ser utilizada para identificar tendências, prever resultados e tomar decisões baseadas em dados. Isso pode levar a decisões mais eficazes e eficientes, bem como a uma melhor comunicação e colaboração entre os diferentes atores envolvidos.

A teoria da TH pode ser uma ferramenta valiosa para promover o desenvolvimento regional. Esta abordagem, ao convergir com a técnica de BI, busca identificar áreas de oportunidade, desenvolver políticas públicas e promover a colaboração e interação entre diferentes atores. Isso pode levar a um aumento na inovação, no desenvolvimento econômico e na qualidade de vida regional.

A teoria da AN configura-se como uma ferramenta essencial no aperfeiçoamento das operações institucionais. Ao empregá-la, visa-se identificar e analisar requisitos cruciais para negócios, especialmente na concepção de novos produtos, caso proposto. Essa teoria, quando integrada à técnica de BI, direciona o uso de abordagens como mineração de dados (em inglês, *Data Mining* – DM) (Apêndice C) e análise preditiva. Tais técnicas possibilitam extrair *insights* valiosos que habilitam as organizações a discernir padrões e oportunidades apresentadas pelos atores da TH e sociedade. Como resultado, as entidades são conduzidas a decisões mais fundamentadas e estratégicas.

A teoria VD objetiva transformar dados brutos em representações gráficas claras, aliando princípios de *design* gráfico, psicologia perceptual e ciência cognitiva. Ao integrar-se com técnicas de BI, tal teoria visa facilitar a identificação de padrões e *insights* em grandes volumes de dados. Essas visualizações mostram-se essenciais no contexto atual da inteligência de negócios, permitindo que os envolvidos na TH e na sociedade busquem soluções já estabelecidas, em progresso ou novas demandas.

Este trabalho objetiva aplicar à RFEPCT em um estudo de caso devido à sua finalidade e objetivo colaborar para a interação TH e promover a inovação. Visa desenvolver um Mínimo Produto Viável ou (em inglês, *Minimum Viable Product* - MVP) na forma de uma plataforma digital para potencializar ecossistemas inovadores. Um MVP é a versão mais básica é um artefato que permite coletar o máximo de aprendizados validados sobre os usuários com o menor esforço possível, cujo objetivo é demonstrar de forma sucinta características com valor claro para o usuário final, sendo um passo crucial na metodologia Lean Startup (Olsen, 2015; Lenarduzzi, et. al., 2016), essencial para testar e validar ideias em ambientes de desenvolvimento, evoluindo de uma versão de novo produto para um conjunto mínimo de funcionalidades.

Tal artefato pretende oferecer *insights* para atender as demandas institucionais, produtivas e governamentais, incluindo aquelas provenientes de órgãos de controle, os quais, quase sempre de forma intempestiva, costumam exigir das ICTs e de seus gestores, dados e relatórios dos mais diferentes tipos, incluindo detalhes que normalmente não figuram no dia a dia da prática da governança dessas instituições. Assim, almeja-se melhorar e contribuir para a evolução do Sistema Nacional de Inovação (SNI), com a aplicação responsável pelo fomento em Ciência e Tecnologia (C&T) e em Pesquisa Desenvolvimento Tecnológico e da Inovação (PD&I), ao corresponder às demandas dos gestores, pesquisadores, setores produtivos e sociedade civil.



## 1.1 Justificativa

A RFEPCT foi selecionada, como estudo de caso (Yin, 2015), considerando sua relevância na promoção do desenvolvimento nacional, por ser o segundo maior ecossistema de Instituições de Ensino público federal brasileiro e contar com o mais alto nível de formação de seus funcionários (docentes e técnicos), ficando atrás apenas das universidades federais, e, por apresentar uma extensa trajetória em colaboração e interação com atores da TH.

De acordo com a Lei nº 11.892/2008 (Basso *et al.*, 2021), o ecossistema da RFEPCT tem como finalidades oferecer educação profissional e tecnológica com ênfase no desenvolvimento socioeconômico local, regional e nacional, desenvolver programas de extensão e divulgação científica e tecnológica, além de estimular a pesquisa aplicada, a produção cultural, o empreendedorismo, o cooperativismo e o desenvolvimento científico e tecnológico.

A RFEPCT também tem como objetivo promover a produção, o desenvolvimento e a transferência de tecnologia social, especialmente aquelas voltadas para a preservação do meio ambiente, com o intuito de ministrar cursos de formação, realizar pesquisas aplicadas e desenvolver atividades de extensão.

Segundo dados públicos disponíveis na plataforma Nilo Peçanha (Brasil, 2018b), até 2018, a RFEPCT contava com uma equipe de 75.996 servidores públicos ativos e efetivos, sendo 11.806 com doutorado e 26.091 com Mestrado, distribuídos em 647 unidades em todos os estados brasileiros. Na época em questão, a RFEPCT atendeu aproximadamente um milhão de estudantes matriculados em todos os níveis de ensino, desde a educação infantil até a pós-graduação. Destaca-se que no ecossistema da RFEPCT o ensino infantil e fundamental, são ofertados exclusivamente pelo Colégio Pedro II (Brasil, 2018b).

Essas características tornam-na uma das maiores ICTs do mundo voltadas ao ensino, o que reforça sua importância na promoção do desenvolvimento do país e na formação de profissionais qualificados em diversas áreas. Portanto, é fundamental investir na melhoria e expansão das ICTs, visando aprimorar ainda mais sua atuação em prol do desenvolvimento socioeconômico do Brasil.

Devido à complexidade inerente ao ecossistema nacional de C&T, é necessário adotar uma abordagem sistêmica que permita compreender, filtrar, gerenciar, disponibilizar e monitorar a evolução dos canais de PD&S em ecossistemas das ICTs. A

dinâmica de inovação é influenciada por múltiplos fatores, incluindo o modelo de inovação TH, que consiste em três atores principais: universidades, indústria e governo (Etzkowitz *et al.*, 2000).

Globalmente, a maioria das ICTs não dispõe das soluções adequadas para gerenciar seus próprios canais de PD&S, mesmo no contexto do modelo TH. No caso brasileiro, essa responsabilidade cabe, especialmente, aos Núcleos de Inovação Tecnológica (NIT) e, internacionalmente, aos seus equivalentes, denominados Escritórios de Transferência Tecnológica (em inglês, TTO - *Technology Transfer Office*).

Uma das atribuições do NIT é promover e captar as demandas relacionadas às interações entre ICTs, SPP e GOV. Além disso, também é responsável por gerir, com foco na comercialização, dentre outras atribuições, os canais de PD&S em ecossistemas das ICTs.

É importante destacar que essas atividades são fundamentais para o processo de inovação, especialmente no campo tecnológico. Portanto, é necessário promover uma abordagem colaborativa e sistemática entre as ICTs, SPP e GOV para impulsionar o avanço da CT&I e PD&I brasileira.

Dentre os métodos mais comuns associados à teoria TH, que busca fomentar a inovação e o desenvolvimento tecnológico entre as ICTs-GOV-SPP, destacam-se, por exemplo a análise de patentes (Basso *et al.*, 2021), publicações científicas (Fan *et al.*, 2015; Rapini *et al.*, 2019), financiamentos (Berbegal-Mirabent *et al.*, 2019; Rapini *et al.*, 2019) e projetos de pesquisa (Berbegal-Mirabent *et al.*, 2019; Rapini *et al.*, 2019). Um método associado a uma teoria é um conjunto de procedimentos ou técnicas sistematizadas, derivadas de princípios teóricos, usadas para aplicar os conceitos dessa teoria na prática. Ele orienta como realizar investigações, análises ou aplicações práticas baseando-se na estrutura conceitual e nos objetivos da teoria correspondente (Gray, 2012).

Estudos de revisão da literatura identificaram também algumas lacunas a serem investigadas considerando o modelo TH. Há necessidade de investigar, por exemplo, alternativas para medir as interações TH de forma mais objetiva (Ankrah *et al.*, 2015) e realizar pesquisas longitudinais (Ankrah *et al.*, 2015; Figueiredo *et al.*, 2020).

Entre as técnicas associadas à TH, a análise fatorial por componentes principais (Sierra *et al.*, 2018) e por agrupamento (Berbegal-Mirabent *et al.*, 2019; Galati *et al.*, 2020) se destacam. No entanto, neste trabalho, também foi observado que a técnica de BI (Johnson *et al.*, 2022) é pouco comum. Mesmo sendo amplamente empregada em

diversas áreas para coletar e analisar informações pertinentes aos negócios (da Cunha *et al.*, 2023; Kongthanasuwan *et al.*, 2023), sua utilização na pesquisa em TH é escassa. Técnicas associadas a uma teoria são procedimentos ou métodos específicos que se baseiam nos princípios e conceitos dessa teoria para a coleta, análise e interpretação de dados, ou para a implementação prática de ideias. Estas técnicas permitem aplicar a teoria a situações reais, facilitando a obtenção de resultados, a resolução de problemas ou a criação de novas abordagens e soluções de acordo com os fundamentos teóricos estabelecidos (Gray, 2012).

A limitada aplicação da BI pode decorrer da falta de familiaridade dos pesquisadores com essa técnica ou da ausência de dados em formatos ideais para tal abordagem. Contudo, o crescente valor dos dados em decisões institucionais e nas ciências sociais indica que a pesquisa em TH utilizando técnicas de BI representa um campo promissor a ser explorado nesta tese.

A integração de dados é um desafio constante, especialmente em bases de dados públicas, conforme descrito em alguns estudos (Barahama *et al.*, 2021; Hoffmann, 2021; Marques *et al.*, 2022). Safadi *et al.* (2016) destacaram que a grande quantidade de dados disponíveis e a sua falta de integração entre as fontes de dados são um obstáculo cada vez maior para o planejamento e um desafio crescente para as agências governamentais. Esse desafio é ainda mais acentuado pela dispersão dos dados em fontes diversas e com padrão limitado de documentação.

A dificuldade na integração dos dados impacta a gestão da inovação, dificultando a tomada de decisões assertivas por parte dos gestores das ICTs, pesquisadores, setores produtivos e sociedade, que buscam implementar políticas e ações, realizar pesquisas, identificar potencialidade de inovações e de fiscalizar execução das políticas públicas, respectivamente embasadas em dados. Grande parte dos dados disponíveis ao público, encontra-se dispersa em diversas fontes, em diferentes formatos e sem padronização. Essa dispersão dificulta ainda mais a integração desses dados. É necessário, portanto, encontrar soluções tecnológicas que mitiguem essa dispersão, com vistas à melhor gestão não só da inovação, mas também dos canais de PD&E das ICTs.

Considerando o período de 2011 a 2018, as ICTs públicas brasileiras foram responsáveis, isoladamente, por aproximadamente 99% de todo conhecimento científico publicado na *Web of Science* (Analytics, 2018; 2019) e 28% da produção tecnológica brasileira, sob a forma de patente (Peixoto *et al.*, 2021) contabilizada no Instituto Nacional da Propriedade Industrial (INPI).

A métrica mais utilizada para medir o desenvolvimento tecnológico no Brasil e no mundo é a quantidade de patentes. No entanto, Suzuki (2018) apresenta uma crítica direta a essa métrica, demonstrando que é possível depositar, publicar e até conceder uma patente seguindo apenas critérios técnicos e administrativos do INPI. Além disso, a falta de clareza nos relatórios de gestão das ICTs dificulta a compreensão, por exemplo sobre qual tipo de patente<sup>2</sup> é considerada, se é de depósito, publicação ou concessão.

O próprio sistema de BI consiste em uma forma de agrupar e explorar informações a partir de dados brutos voltados para empresas específicas. O governo brasileiro utiliza duas grandes plataformas de BI para acessar dados das ICTs, as plataformas Nilo Peçanha (Brasil, 2018b) e Painel Universidade 360 (Brasil, 2021a), mantidas pelo governo federal para monitor e avaliar a estrutura administrativa, financeira e de ensino do ecossistema das ICTs públicas federais, com base em dados e para auxiliar a tomada de decisão governamental.

Com base na literatura sobre o assunto, pode-se argumentar que a BI é uma técnica que engloba um conjunto de teorias, metodologias, processos e tecnologias cujo principal objetivo é transformar dados brutos em informações significativas e úteis para a tomada de decisões institucionais (Bimonte *et al.*, 2021). Essa transformação é facilitada pela criação de DW (Zhang *et al.*, 2022) e é alimentada por processos de extração, transformação e carregamento de dados (em inglês, ETL) (Bimonte *et al.*, 2021; Fana *et al.*, 2021).

A BI também abrange processos como Processamento Analítico *Online* (em inglês, *Online Analytical Processing* OLAP) e *data mining* (DM), bem como tecnologias como *dashboard*. DW, conforme definido por Ammar *et al.* (2022), é um banco de dados multidimensional otimizado para consultas analíticas provenientes de diferentes fontes.

A técnica de BI é muito utilizada nas instituições do SPP, pois permite que elas obtenham informações relevantes a partir de seus dados. No entanto, sua implementação exige uma compreensão sólida dos processos envolvidos e a presença de profissionais

---

<sup>2</sup> Depósito, publicação e concessão são etapas distintas, frequentemente associadas ao processo de obtenção de patentes. O "depósito" refere-se à apresentação inicial de um pedido de patente junto a um órgão competente, que dá início ao processo de análise. A "publicação" ocorre quando os detalhes da invenção são tornados públicos, geralmente após um período determinado desde o depósito, permitindo que terceiros tenham conhecimento e possam contestar a originalidade ou a validade da invenção. Já a "concessão" é a etapa final em que, após avaliação e cumpridos todos os requisitos, a patente é efetivamente concedida ao requerente, garantindo-lhe direitos exclusivos sobre a invenção por um período específico, sendo essa última a mais relevante e almejada de se obter.

capacitados para desenvolver soluções gerenciáveis e interpretáveis pelo público em geral, especialmente os tomadores de decisão.

Os gestores das ICTs, especialmente as públicas, têm a responsabilidade de cumprir os requisitos da Lei nº 13.243/2016 e de entidades como o Tribunal de Contas da União (TCU), Controladoria Geral da União (CGU), Ministério da Educação (MEC), CAPES, SETC, SESU e o Ministério da Ciência e Tecnologia (MCTI), que inclui a geração de relatórios conformes. Simultaneamente, é essencial que adotem tecnologias avançadas para melhorar a gestão dos canais de PD&S, permitindo um planejamento estratégico mais assertivo, implementação de melhores práticas e tomadas de decisão baseadas em dados, considerando os princípios norteadores da Lei Geral de Proteção de Dados Pessoais – Lei nº 13.709/2018 (LGPD). Essa abordagem não só atende às exigências regulatórias, mas também potencializa a inovação e o desenvolvimento socioeconômico regional.

## **1.2 Problema**

Diante da crescente importância do atendimento às demandas institucionais, produtivas e governamentais no ecossistema de Ciência, Tecnologia e Inovação, especialmente aquelas demandas oriundas de órgãos de controle, sociedade e setores produtivos surge um problema crítico: como a gestão dos dados de pesquisa, desenvolvimento tecnológico e prestação de serviços (PD&S) nas Instituições de Ciência e Tecnologia (ICTs) pode ser organizada e aprimorada para responder de maneira tempestiva e confiável a essas exigências?

## **1.3 Questão de Pesquisa**

Uma solução em BI, embasada nas teorias de Análise de Negócio, Visualização de Dados e Tripla Hélice de Inovação e avaliada por especialistas, é um instrumento tecnológico necessário para que a sociedade possa aprimorar a análises e a identificar os impactos dos canais de PD&S das ICTs no desenvolvimento regional. Isso é feito por meio de mapeamento e análise dos dados, contribuindo para a criação de políticas públicas eficazes, monitoramento de avanços, estímulo à inovação e à colaboração, promovendo transparência e desenvolvimento socioeconômico e atendimento as exigências legais e administrativas do governo, por meio dos seus órgãos de controle.

## 1.4 Tema

O tema do presente trabalho visa a integração e gestão de dados dos canais de PD&S em ICTs com uma Abordagem Baseada em BI e sustentada pelas teorias de Análise de Negócio, Visualização de Dados e da Tripla Hélice de Inovação para otimização da tomada de decisão e desenvolvimento regional.

## 1.5 Tese

Postula-se que, ao conceber e implementar soluções tecnológicas em BI, ancoradas nas teorias de Análise de Negócios, Visualização de Dados e da teoria da Tripla Hélice de Inovação, permite às ICTs e a sociedade gerar *insights* a partir do dados organizados utilizando recursos de visualização de dados e, ao mesmo tempo, contribuir para que seus gestores possam responder de forma tempestiva e confiável às crescentes demandas institucionais, produtivas e governamentais, as quais inclui as demandas provenientes de órgãos de controle, da sociedade e dos setores produtivos, a partir dos dados vinculados aos canais de PD&S em ICTs.

## 1.6 Hipótese

- $H_a$ : A solução tecnológica proposta é relevante e inovadora no apoio à tomada de decisão, análise, gestão e monitoramento baseados em evidências dos canais de PD&S em ecossistema das ICTs, auxiliando os gestores a atenderem as exigências legais e administrativas do GOV e as demandas por soluções pelos SPP.

## 1.7 Objetivo Geral

Integrar e otimizar a gestão dos dados públicos referentes aos canais de Pesquisa, Desenvolvimento Tecnológico e Serviços (PD&S) das Instituições Científicas, Tecnológicas e de Inovação (ICTs) e contribuir, por meio da solução tecnológica desenvolvida neste trabalho, para melhorar a transparência e responder de maneira eficaz e confiável às demandas institucionais, produtivas, governamentais e legais, promovendo o desenvolvimento socioeconômico e a inovação no contexto regional.

## 1.8 Objetivos Específicos

- Identificar as fontes de dados relevantes e necessárias para a implementação da solução proposta;
- Desenvolver um modelo de dados adequado que permita a obtenção de informações precisas e objetivas a partir das fontes de dados identificadas;
- Projetar e desenvolver um sistema capaz de integrar e analisar as diferentes fontes de dados, utilizando técnicas de BI, DW e um painel interativo sob a forma de um *dashboard*, de maneira objetiva e confiável;
- Validar por meio de pareceres técnicos fornecidos por especialistas envolvidos em ecossistema das ICTs se a integração e a otimização dos dados e a solução tecnológica desenvolvida neste trabalho é considerada inovadora contribui para a solução do problema proposto;
- Validar se os dados exportados pela solução tecnológica daqui desenvolvida, após a seleção de um determinado cenário, considerando os dados do estudo de caso selecionado, permite desenvolver uma análise de agrupamento e a criação de um índice integrado de performance e inovação em PD&S (Apêndice C);

## 1.9 Contribuições Esperadas

- **Integração teórica robusta:** propor uma integração de três corpos teóricos, ou seja, a Análise de Negócio, a Visualização de Dados e a Tripla Hélice de Inovação. Ao aplicá-los ao contexto das ICTs e à gestão de dados dos canais de PD&S, pode-se esperar uma abordagem inovadora e interdisciplinar, com vistas às questões da gestão de dados em organizações complexas.
- **Solução avançada de BI:** desenvolver uma solução tecnológica *online* e pública em BI que se aprofunde não apenas na visualização, mas também na análise de dados multivariados, representa uma abordagem avançada. Ao usar tecnologias de *Data Warehouse* e *Dashboard*, a solução tecnológica pode fornecer *insights* sobre como melhorar a extração, tratamento, carregamento, organização e exploração de dados brutos em contextos institucionais e governamentais.

- **Metodologia para integração de dados:** identificar fontes de dados, desenvolver um modelo de dados e criar um sistema para integrar e analisar várias fontes. Esta metodologia sistemática pode oferecer um *framework* padronizado.
- **Impacto na tomada de decisão e gestão:** avaliar o impacto da solução proposta na tomada de decisão e na gestão das ICTs, revelar práticas inovadoras e *insights* sobre como as soluções tecnológicas *online* e públicas influenciam a responsividade e o cumprimento das ICTs ao problema supracitado.
- **Validação prática:** validar a solução proposta, por meio de estudos de caso ou pareceres técnicos fornecidos por especialistas, para proporcionar *insights* reais sobre sua aplicabilidade. Essa validação empírica não apenas contribui para a robustez deste estudo, mas também fornece um caminho para a implementação prática de suas recomendações em contextos semelhantes.

Diante da complexa interação entre os diferentes atores envolvidos no desenvolvimento regional, a proposição, gestão e o monitoramento dos canais de PD&S em ecossistema das ICTs, que utilizam como referência o modelo TH, representa um grande desafio.

Ainda que haja outras soluções para outros contextos, a ausência de uma solução tecnológica *online*, pública, confiável para a obtenção de dados e informações precisas e objetivas, tem dificultado a tomada de decisão, análise e gestão embasadas em evidências, considerando o escopo apresentado neste trabalho.

Entende-se que a adoção da solução tecnológica proposta poderá contribuir para o avanço do conhecimento dos canais de PD&S em ecossistema das ICTs, oferecendo subsídios práticos à tomada de decisão, investigação, gestão e monitoramento embasados em evidências.

Este trabalho está organizado da seguinte forma: Seção 2 apresenta a fundamentação teórica, seguido pela Seção 3, que descreve a metodologia utilizada. Os resultados são apresentados na Seção 4 e, na Seção 5, as análises e discussões. Por fim, na Seção 6 são apresentadas as conclusões e, na Seção 7 propostas de pesquisa e desenvolvimento tecnológicos futuros. Os apêndices servem de suporte aos detalhes técnicos necessários para uma melhor compreensão dos procedimentos realizados.



## 2 Referencial Teórico

A interação e/ou a colaboração entre as ICTs com os SPP são consideradas elementos cruciais para o desenvolvimento econômico, regional e à promoção da inovação tecnológica. A compreensão dessas relações é fundamental para o planejamento e a implementação, de um lado, de políticas públicas e, de outro, de estratégias institucionais.

Para avaliar a natureza e o impacto dessas relações, muitos estudiosos, por exemplo Leydesdorff et al. (1996; 1998), Lino (2013), Silva (2015), Mould et al. (2016), Jackson et al. (2018), Menezes et al. (2019), Yaakub et al. (2020), Barragan-Ocana et al. (2020), Stewart et al. (2022) têm apresentado teorias e métodos para medir a interação entre ICTs e o setor produtivo. Através dos seus canais de PD&S, elas desempenham papéis importantes na sociedade. Elas são responsáveis por formar profissionais qualificados, por desenvolverem ciência e tecnologia, buscando soluções e/ou prestação de serviços científicas e/ou tecnológicas à sociedade, ao GOV e ao SPP.

As ICTs são responsáveis, também, por desenvolverem pesquisas ou prestação de serviços científicos e/ou tecnológicos, cujo objetivo é transferir conhecimento e tecnologia para o SPP, a exemplo da Empresa Brasileira de Pesquisa Agropecuária (EMBRAPA) e do Instituto de Pesquisas Tecnológicas (IPT).

O SPP é composto pelas empresas que produzem bens e serviços para o mercado, incluindo tanto as empresas privadas quanto as empresas públicas, a exemplo da Petrobras e da Financiadora de Estudos e Projetos (Finep). Este setor é fundamental para o desenvolvimento econômico, uma vez que é responsável pela geração de emprego e renda e pelo financiamento e promoção da inovação através de suas demandas diretas ou indiretas às ICTs.

As instituições que integram o SPP compartilham objetivos comuns, incluindo a maximização da eficiência, a promoção da inovação e o crescimento econômico. Por esse motivo, esse setor pode ser identificado pela sigla SPP, que representa a integração de todas as instituições produtivas, incluindo tanto as empresas privadas quanto as empresas públicas.

Na literatura há diversas teorias, métodos, técnicas e ferramentas que foram propostas para investigar os diversos tipos de interação entre as ICTs e os SPP, com o objetivo de avaliar a natureza e o impacto dessas relações. A seguir são apresentadas as

teorias, os métodos, as técnicas e as ferramentas utilizadas para investigação nesta área do conhecimento.

Dentre os diversos estudos consultados, diversas teorias foram utilizadas para compreender os fenômenos relacionados aos diversos tipos de interações entre as ICTs e os SPP, as quais são listadas a seguir:

- **Teoria da Triple Hélice - TH** (em inglês, *Triple Helix Theory*): enfatiza a importância da interação entre ICTs, SPP e GOV na promoção da inovação e desenvolvimento econômico, social e tecnológico de uma região ou país. Esta teoria destaca a colaboração entre esses três atores como fundamental para impulsionar a inovação e o desenvolvimento regional ou nacional (Etzkowitz *et al.*, 2000).
- **Teoria do Sistema Nacional de Inovação - SNI** (em inglês, *National Systems of Innovation Theory*): sugere que a inovação é resultado da interação entre diferentes setores, incluindo ICTs, SPP e GOV e outras instituições. Esta teoria destaca a importância do ambiente institucional e infraestrutural de um país na promoção da inovação (Freeman, 1987).
- **Teoria da Inovação Aberta** (em inglês, *Open Innovation Theory*): propõe que os SPP devem olhar além de suas fronteiras internas, aproveitando ideias externas e colaborações para impulsionar a inovação. Esta teoria desafia o modelo tradicional de P&D fechado, promovendo uma abordagem mais permeável e colaborativa (Chesbrough, 2003).
- **Teoria da Gestão do Conhecimento - GC** (em inglês, *Knowledge Management Theory*): foca nos processos organizacionais relacionados à criação, compartilhamento, uso e gestão do conhecimento. Esta teoria vê o conhecimento como um ativo valioso que, quando gerenciado corretamente, pode levar a vantagens competitivas e inovação (Drucker, 1993).
- **Teoria da Inovação Sistêmica** (em inglês, *Systems of Innovation Theory*): enquanto semelhante à Teoria do SNI, esta teoria tem uma abordagem mais ampla, considerando as redes e interações entre diferentes atores (não apenas no nível nacional) que contribuem para a inovação. Ela examina o sistema mais amplo e as interconexões que facilitam a inovação (Breschi *et al.*, 1997).

As teorias listadas acima são amplamente utilizadas na literatura científica para compreender as diversas interações que ocorrem entre ICTs, SPP e GOV cujo objetivo a ser alcançado é o desenvolvimento de uma região, preferencialmente pela inovação. Com características aparentemente redundantes, oferecem uma perspectiva diferente sobre o papel de cada um dos três atores envolvidos e outras características (Apêndice B, Quadro B. 2, Quadro B. 2).

Considerando que a teoria TH é amplamente utilizada na literatura científica, especialmente a partir do ano 2000; que apresenta uma abordagem interdisciplinar; que considera a não sobreposição dos atores envolvidos; e que os propósitos e objetivos norteadores da RFEPCT estão em consonância com o modelo TH, ela será utilizada neste trabalho como base teórica, com vistas à compreensão do estudo de caso apresentado (Yin, 2015), à interação entre a RFEPCT e SPP e à contribuição à promoção da inovação e desenvolvimento econômico, social e tecnológico de uma região ou país.

Após a definição da teoria TH como suporte teórico para explicar o fenômeno observado, são apresentados alguns dos métodos mais comuns encontrados durante a revisão da literatura, associados à TH, tais como a análise das: i) redes sociais (Paswan *et al.*, 2022); ii) patentes (Basso *et al.*, 2021); iii) transferência de tecnologia (Rapini *et al.*, 2019); iv) publicações científicas (Fan *et al.*, 2015; Rapini *et al.*, 2019); v) financiamentos (Berbegal-Mirabent *et al.*, 2019; Rapini *et al.*, 2019); vi) projetos de pesquisa (Berbegal-Mirabent *et al.*, 2019; Sierra *et al.*, 2018); vii) *spin-offs* (Galati *et al.*, 2020); viii) licenças de tecnologia (Tseng *et al.*, 2020); ix) quantitativa (Bernardino *et al.*, 2020; Kang *et al.*, 2019) e x) qualitativa (Bernardino *et al.*, 2020; Purbasari *et al.*, 2020).

Considerando que o método quantitativo permite incorporar em sua análise um conjunto de múltiplos indicadores de saída, ele é o mais utilizado para analisar estudos fundamentados na teoria TH. É amplamente utilizado na engenharia industrial para avaliar e solucionar problemas relacionados à produção. Esse método baseia-se em dados quantitativos e análise estatística para compreender e modelar o comportamento de sistemas de produção das ICTs. Permite uma avaliação objetiva e eficiente dos resultados de produção e apresenta soluções otimizadas, razão pela qual foi selecionado para analisar os resultados da RFEPCT. A partir das considerações apresentadas, adotou-se para esta tese o método quantitativo.

Após a definição da teoria e do método, são apresentadas algumas das técnicas mais comuns relacionadas à análise multivariada de dados encontradas durante a revisão da literatura, associadas à TH; são elas, por exemplo: i) análise fatorial por componentes

principais (Sierra *et al.*, 2018); ii) *business intelligence* (Johnson *et al.*, 2022); iii) análise de agrupamento (Berbegal-Mirabent *et al.*, 2019; Galati *et al.*, 2020); iv) análise de correspondência múltipla (Albats *et al.*, 2022); v) análise de regressão (Berbegal-Mirabent *et al.*, 2019; Cartaxo *et al.*, 2017); vi) análise de variância (Baglieri *et al.*, 2018; Galati *et al.*, 2020); vii) análise de covariância (Leydesdorff *et al.*, 2009); viii) árvore de decisão (Laursen *et al.*, 2011); e x) séries temporais (Khatibi *et al.*, 2017; Souzanchi Kashani *et al.*, 2018).

Conforme também constatado, há uma quantidade de técnicas multivariada utilizadas para auxiliar na análise dos dados a partir da teoria TH. Identificou-se que as técnicas mais utilizadas são aquelas comumente aplicadas à análise estatística.

Esta pesquisa busca contribuir com essa área de estudo, uma vez que apenas em um trabalho, até o momento, baseou-se na teoria da TH, a aplicação da técnica de *business intelligence* (BI). No Brasil, embora existam exemplos de soluções tecnológicas de acesso público ou particulares similares (Apêndice B, Quadro B. 1), até o presente momento segundo alguns especialistas, nenhuma contém as soluções propostas neste MVP (Subseção 4.4 e Apêndice B, Quadro B. 3). A contribuição para essa área de estudo deve-se ao fato de que a técnica de BI permite coletar, armazenar, analisar e apresentar grandes quantidades de dados para auxiliar na tomada de decisões informadas.

Dentre as tecnologias disponíveis para o desenvolvimento de soluções de BI, é possível destacar *data mining*, análise de dados, *dashboards* interativos e relatórios, que ajudam a visualizar e compreender dados complexos, identificar tendências, padrões e tomar decisões baseada em dados.

Estudos recentes sobre a TH, identificaram que: i) há necessidade de investigar alternativas para medir de forma mais objetiva as interações TH, além da medida subjetiva atualmente empregada (Ankrah *et al.*, 2015); ii) a maioria dos estudos revisados são estudos transversais, havendo necessidade de pesquisas longitudinais (Ankrah *et al.*, 2015; Figueiredo *et al.*, 2020); iii) subjetividade na escolha de palavras-chave e uso de apenas um banco de dados (Ankrah *et al.*, 2015; Figueiredo *et al.*, 2020); e iv) conhecimento dos canais usados nas interações entre os gestores (Galvão *et al.*, 2019).

A técnica escolhida foi a BI associada a aplicações das tecnologias de DW, *dashboard* interativo e análise de dados. Esse conjunto de técnica e tecnologias foi empregado para propor uma solução baseada em um produto digital (*software*) que possibilite obter conhecimentos precisos e confiáveis sobre os canais de PD&S em ICTs.

Contudo, para que uma técnica de BI seja efetivamente aplicada e gere *insights* valiosos, é fundamental que esteja ancorada nas teorias relacionadas à área (Seção 2.4.2).

O objetivo do MVP proposto é possibilitar aos usuários a tomada de decisões baseadas em evidências para mitigar as avaliações subjetivas. Ele será usado para demonstrar a possibilidade e o potencial da técnica de BI, contribuindo para o aprimoramento da tomada de decisão (Santos, O. *et al.*, 2019).

A abordagem MVP é uma técnica amplamente utilizada para desenvolver produtos que são bem-sucedidos (Ries, 2011). Ao adotá-lo, os desenvolvedores podem aprender rapidamente o que os usuários querem e precisam, evitar o desperdício de tempo e recursos e tomar decisões baseadas em dados sobre como melhorar seus produtos (Liao *et al.*, 2019).

No entanto, é importante notar que os dados representam apenas o que se tem e não necessariamente a realidade das ICTs, especificamente aquelas vinculadas à RFEPCT; também, que o *dashboard* interativo disponibiliza a visualização dos dados disponíveis e, a partir das tabelas de dado geradas pelos filtros, realizar análises multivariadas. Embora os dados sejam “*falibilísticos*”, é provável que mais dados apresentem outras informações (Macqueen, J, 1967; Santos, E. *et al.*, 2019).

Após a definição da teoria, do método e da técnica e das tecnologias, são apresentadas algumas das ferramentas mais comuns encontradas durante a revisão da literatura, associados à técnica de BI para o desenvolvimento do produto, quais sejam: i) Tableau; ii) Power BI; iii) SAP Lumira; iv) IBM Cognos; v) *QlikView*; vi) *Oracle Business Intelligence*; vii) *MicroStrategy*; viii) *TIBCO Spotfire*; ix) *SAS Business Intelligence*; e x) *Oracle Data Visualization*.

A escolha sobre qual melhor ferramenta de BI deve ser utilizada depende de alguns fatores, como os requisitos dos usuários e de análise de dados, a integração com outras ferramentas e o orçamento disponível.

Considerando, dentre outras características, que: o Power BI permite criar visualizações de dados interativos, exportações dos dados filtrados em formato de tabelas e compartilhá-las com outras pessoas em uma plataforma baseada na nuvem; oferece recursos avançados de análise de dados, como *dashboards* e relatórios, além de integração com outras ferramentas da *Microsoft* como o *Excel*; é de fácil aprendizagem; disponibiliza grande comunidade de suporte *online* e, finalmente, para a academia, sua licença de uso é gratuita, não havendo restrições de funcionalidades avançadas, o Power BI foi utilizado

neste trabalho como a ferramenta de desenvolvimento do MVP proposto, utilizando a linguagem DAX da *Microsoft*.

A seguir, será apresentada uma revisão dos canais de PD&S em ICTs vinculadas à RFEPCT. Na sequência, uma revisão das condições da infraestrutura nacional de PD&I e CT&I brasileira durante a primeira década de vigência do ecossistema da RFEPCT. Logo após, uma revisão da evolução da teoria TH de inovação. Por fim, será apresentada uma revisão da literatura sobre BI. Essas revisões foram realizadas com o objetivo de fornecer uma base teórica para a discussão dos resultados da pesquisa realizada no âmbito desse trabalho.

## **2.1 Revisão dos Resultados Relacionados aos Canais de Pesquisa, Desenvolvimento Tecnológico e Prestação de Serviços da RFEPCT**

O mapeamento dos dados referentes aos canais de PD&S em ecossistema das ICTs permite monitorar e planejar os diversos desenvolvimento, por exemplo, o regional por meio da colaboração entre gestores, pesquisadores e sociedade.

O canal de pesquisa refere-se as atividades de pesquisa são conduzidas, divulgadas e acessadas. Estes canais podem incluir, mas não se limitam a publicações acadêmicas, bancos de dados de pesquisa, repositórios institucionais, conferências científicas, além de plataformas online e digitais que facilitam a colaboração, o compartilhamento de conhecimento, o acesso a recursos de pesquisa e a comunicação entre pesquisadores. O objetivo desse canal é disseminação de descobertas, a colaboração interdisciplinar e a contribuição para o avanço do conhecimento e da inovação tecnológica.

O canal de desenvolvimento tecnológico refere-se as atividades de pesquisa aplicada, desenvolvimento de propriedades industriais (licenciamento de tecnologia, patentes, registros de softwares, etc.), programas de financiamento para P&D, bem como publicações técnicas e grupos de pesquisa especializados que promovem a disseminação do conhecimento tecnológico. O objetivo desse canal é acelerar a transferência de conhecimento do ambiente acadêmico para o mercado, fomentar a inovação, impulsionar o desenvolvimento econômico e atender às necessidades sociais por meio de soluções tecnológicas avançadas.

O canal de prestação de serviços refere-se aos meios pelos quais essas instituições oferecem serviços baseados em conhecimento e tecnologia ao setor público, privado e à sociedade em geral. Estes serviços podem incluir consultorias, análises técnicas, testes de

laboratório, formação profissional e educacional, transferência de tecnologia, apoio à inovação em empresas, entre outros. O objetivo desses canais é aplicar o conhecimento científico e as capacidades tecnológicas desenvolvidas nas ICTs para solucionar problemas práticos, contribuir para o desenvolvimento econômico, fomentar a inovação, promover a inclusão social e atender às demandas da sociedade e do mercado.

O monitoramento e planejamento desses canais é importante porque pode: i) fomentar o avanço inovativo; ii) compreender objetivamente as demandas técnico-científicas; iii) auxiliar na elaboração de políticas públicas; iv) possibilitar ações estruturantes baseadas no estudo de variáveis inerentes à propriedade intelectual científica e tecnológica; v) tornar os serviços técnicos prestados pelas ICTs mais eficientes na solução de demandas reais do governo, dos pesquisadores, dos setores produtivos e da sociedade em geral; e vi) assistir os gestores no atendimento às demandas por dados e informações dos canais de PD&S aos órgãos de controle.

O mapeamento de metadados é uma ferramenta importante para as ICTs atenderem aos seus objetivos e finalidades. Portanto, é recomendável que seus gestores invistam em ferramentas que apresentem esse mapeamento o mais detalhadamente possível. Esse mapeamento também permitirá que a sociedade possa compreender a evolução dos canais de PD&S em ICTs ao longo do tempo. Com essa compreensão, os gestores, em especial, poderão estabelecer ações e políticas institucionais.

Apresenta-se, em seguida, uma revisão da literatura, com o objetivo de investigar quais dados referentes aos canais de PD&S do ecossistema da RFEPCT já foram estudados desde sua criação em 2008.

Garcia (2011), ao investigar o desenvolvimento da produção científica e tecnológica no Instituto Federal da Paraíba (IFPB), utilizou como principal base de dados a plataforma Lattes e o Diretório do Grupo de Pesquisa (DGP). Ele analisou a produção dos docentes pesquisadores, entre 1994 e 2008 e constatou uma diferença aproximada de 4 vezes entre a quantidade de publicações científicas (total = 2.788), quando comparada às publicações tecnológicas (total = 512). Ele ainda reconheceu que, devido as finalidades e objetivos históricos dos CEFET, era esperado um resultado “invertido” e estabeleceu a necessidade de construções de indicadores da produção científica e tecnológica, pois eles contribuem para que os gestores definam políticas institucionais mais aderentes às finalidades e objetivos das instituições.

Lino (2013), após um estudo qualitativo da política da Gestão do Conhecimento (GC) adotada no IFC, identificou a necessidade de elaboração de diretrizes para sua

implementação nos Institutos Federais (IF). Conclui seu estudo destacando que o Instituto Federal de Santa Catarina (IFSC) institucionalizou a GC em 2004, portanto ainda CEFET, e essa institucionalização ocorreu de forma sistêmica e experimental. Ela ainda destacou que, no início dos anos 2000, não havia outra ICT no ecossistema dos CEFET com uma proposta similar. Ela também constatou que poucas ICTs públicas tinham implementado alguma ação formalizada nesse sentido e que, à época, não existia literatura adequada que pudesse ser usada como suporte para a implementação da GC em ecossistema das ICTs públicas. Ao final, ela apresentou alguns *gaps* dentre os quais a demanda por desenvolver um “estudo quantitativo para a verificação e sistematização das instituições de ensino que adotam a gestão do conhecimento e, principalmente, quais são os passos e/ou modelos seguidos” (Lino, 2013).

Niwa (2014) analisou a aplicação do modelo TH, tendo como pano de fundo as incubadoras tecnológicas instaladas na Universidade Tecnológica Federal do Paraná (UTFPR). Ela constatou que o então modelo administrativo das ICTs, no Brasil, apresentaram vícios que fragilizam sua interação no cenário da TH. Segundo essa autora, existem três fatores que impedem a plena interação, são eles o excesso de burocracia, a lentidão nos procedimentos e a desmotivação dos docentes e técnicos.

Macedo (2014), ao analisar as características dos grupos de pesquisa do Instituto Federal da Bahia (IFBA), entre 2008 e 2013, disponíveis na fonte de dado do DGP, constatou indícios de que os grupos não se consolidaram, uma vez que seus membros aparentemente não produzem conhecimento compartilhado internamente na ICT. Ao final de sua análise, a autora apresentou alguns *gaps* para estudos futuros, como a inclusão de outros indicadores, por exemplo, patentes ou outra variável que possa captar a interação entre ICTs e SPP, através dos seus grupos de pesquisa.

Santos *et al.* (2015) analisaram a produção científica do Instituto Federal de São Paulo (IFSP). A fonte de dados utilizada foi a plataforma Lattes. Os currículos foram coletados em novembro de 2013 e o estudo analisou os indicadores da produção bibliográfica entre os anos de 2009 e 2011. Os indicadores considerados foram: a) artigos completos publicados em periódicos; b) trabalhos completos publicados em anais de congressos; c) resumos expandidos publicados em anais de congressos; d) resumos publicados em anais de congressos; e e) apresentações de trabalho. Eles selecionaram 427 docentes que atuavam no ensino superior para realizar seu estudo. Desse total, 33 (7%) não possuíam cadastro na plataforma Lattes. Eles constataram que, apesar de uma tendência de crescimento de 21,9% ter sido observada na produção científica, analisada



entre 2009 e 2011, constataram que antes e depois da mudança do CEFET/SP para o IFSP, não houve alterações nessas produções. Ao concluírem sua análise, eles apresentam duas ressalvas relacionadas ao possível crescimento, ou seja, que as publicações poderiam estar vinculadas a outras instituições e que poderia ter havido um aumento na oferta de bolsas de iniciação científica no período.

Perucchi *et al.* (2015) analisaram a produção da pesquisa de 96 professores de uma amostragem aleatória estratificada em 38 IF, com o objetivo de investigar como se estabeleceu a produção científica do ecossistema da RFEPCT, considerando os estímulos, apoio e divulgação fomentados pelos seus gestores, entre os anos de 2009 e 2014. A fonte de dados utilizada foi um questionário coletado neste último ano. Eles constataram que, dos 18,71% da produção de pesquisa resultante, 15,48% foram publicações de artigos científicos completos e 3,23% publicações referentes à inovação. Eles destacam que, dentre o percentual de inovações publicadas, não foi identificado nenhum registro de patentes. Além disso, constataram também que 77% dos entrevistados relataram que sua motivação para desenvolver pesquisas não estava relacionada às demandas do setor produtivo, mas sim aos seus objetivos acadêmicos. Eles concluíram enfatizando que os entrevistados também relataram ter recebido pouco ou nenhum apoio institucional de seus gestores para desenvolverem suas atividades de pesquisa.

Silva (2015) focou o estudo na identificação de possíveis fatores propulsores dos canais de PD&S do ecossistema da RFEPCT. Foram considerados os Institutos Federais de Santa Catarina (IFC), de Goiás (IFGO), do Espírito Santo (IFES), Catarinense (IFC), do Rio Grande do Sul (IFRS), do Ceará (IFCE), Goiano (IFGOIANO) e Instituto Federal do Sul de Minas Gerais (IFMG). As fontes utilizadas referem a duas chamadas, quer sejam a do CNPq-SETEC/MEC nº 94/2013 e a do CNPq-SETEC/MEC nº 17/2014, INEP, DGP. Também foram utilizados, no supracitado estudo, Indicadores de Gestão das Instituições Federais da Rede EPCT/MEC/TCU, por exemplos índice dos cursos de graduação e pós-graduação, índice de titulação do corpo docente e diretório de grupos de pesquisa. Ela apresenta as seguintes constatações: i) IFSC, IFGO, IFES, IFCE e IFRS apresentaram, na ordem, o melhor desempenho em relação aos demais participantes que concorreram na linha de apoio de PD&I; ii) os que apresentaram melhor desempenho na linha de apoio de extensão tecnológica foram, pela ordem IFES, IFSC, IFRS, IFGOIANO, IFSMG e IFCE; iii) “cinco Institutos estão relacionados nas duas chamadas como os de melhor desempenho, a saber: IFSC, IFGO, IFES, IFCE e IFRS; iv) o Instituto Federal Rio Grandense (IFSUL) e o Instituto Federal do Sul de Minas (IFSMG) estão

relacionados em uma das duas chamadas; v) que os cinco IF “com melhor desempenho nas duas Chamadas Públicas” são os que apresentam “os maiores quantitativos de Grupos de Pesquisa, considerando todas as Instituições que integram a RFEPCT” e vi) que dos sete IF que apresentam “o melhor desempenho nas duas Chamadas Públicas”, seis estão dentre os oito melhores Instituições de Gestão do Conhecimento (IGCs) da RFEPCT. A exceção foi o IFCE com IGC/2013 igual a 3, ficando o IFSC em 2º lugar.

Perucchi *et al.* (2017) analisaram a produção científica, tecnológica e de prestação de serviços referente a 38 IF. A base utilizada foi a plataforma Lattes, sendo que tais currículos foram coletados entre fevereiro e março de 2014. O estudo analisou os indicadores entre os anos de 2009 e 2014. Eles identificaram, à época, uma população de 24.335 docentes efetivos e a amostra utilizada foi de 165. Eles não informaram quanto da população não possuía cadastro na plataforma Lattes. Eles consideraram apenas os currículos Lattes atualizados entre 2013 e 2014 e com produção cadastrada. Eles concluíram que 88 (53%) dos pesquisados apresentaram alguma produção bibliográfica e dentre essa produção, apenas 17% eram artigos; 45 (27,2%) registraram alguma produção técnica. Dentre aqueles que informaram algum tipo de prestação de serviço, não foram encontrados registros referentes à consultoria. Eles constatam, também, o “baixo número de ocorrências para quase todos os itens de produção técnica”, sendo que 50 (30,3%) dos docentes apresentaram alguma produção em desenvolvimento tecnológico. Dentre essa produção, não foram encontrados registros referentes a patentes. Eles ao concluir sua análise, apresentaram ressalvas relacionadas aos resultados, a saber: i) que a baixa prestação de serviços pode estar relacionada a “produtos ou processos que requerem sigilo”; ii) “a existência de currículos Lattes atualizados sem registros de produção em nenhum dos grupos analisados”; iii) a baixa prestação de serviços e de desenvolvimento tecnológico em contraste com quantidade de produção bibliográfica; iv) “a falta de clareza das terminologias adotada pelo CNPq”; v) “o preenchimento do currículos Lattes nem sempre é corretamente realizado”; vi) os professores/pesquisadores dos Institutos Federais “dão preferência à comunicação informal”; vii) “a produção de artigos científicos formalmente referendados também pode ser considerada significativa, pois é mais frequente que a modalidade das produções técnicas e de inovação”.

Bento (2015) analisou a produção científica do Centro Federal do Rio de Janeiro (CEFET/RJ), Centro Federal do de Minas Gerais (CEFET/MG), Instituto Federal do Rio de Janeiro (IFRJ), Instituto Federal Fluminense (IFF), UTFPR e CPII, relacionada aos seus periódicos institucionais. A principal base utilizada foi a plataforma Lattes. O autor

não deixou claro qual o período considerado. Ele constatou que: i) “as instituições encontram dificuldades para manter as suas publicações correntes” e ii) “a não compreensão da diferença entre os regimes de informação resultantes de experiências distintas, como formação no ensino médio e formação no ensino superior”.

Miano; Do Couto; *et al.* (2017) mensuraram o desenvolvimento da pesquisa no CEFET/RJ, IFRJ e IFF. Eles apresentaram um conjunto de critérios para excluir o CPII da análise. A principal base utilizada por eles foi o DGP. Analisaram os grupos de pesquisa criados entre 2008 até 2016. Foi constatado, que: i) as pesquisas foram desenvolvidas por pesquisadores com diversas formações, não sendo exclusivo por doutores; ii) houve crescimento de grupos de pesquisa e iii) houve uma concentração dos grupos nas grandes áreas, na ordem, grupos nas Ciência Exatas e da Terra (total = 37), nas Ciências Humanas (total = 30), nas Engenharias (total = 28) e nas Ciências da Saúde (total = 16). Ainda segundo eles, “os dados obtidos vão de encontro a outros levantamentos sobre a produção acadêmica e mobilização de grupos de pesquisa dos IF” (Miano; Do Couto; *et al.*, 2017).

Chaves *et al.* (2017) analisaram o desenvolvimento da produção científica e tecnológica no IF Baiano. As principais bases utilizadas foram o DGP e o INPI. Os dados foram coletados em 2017 e foi analisada a produção entre 2010 até 2016. Eles constataram que: i) “submissões de pedido de patente passaram a ser realizados a partir de 2015”; e ii) houve uma concentração dos grupos nas grandes áreas, na ordem, Ciências Humanas (com total = 6), Ciências Agrárias (total = 3), Ciências Exatas e da Terra (total = 2) e Ciências Biológicas (total =1).

Miranda *et al.* (2018) propuseram agrupar os IF considerando suas características como gastos com investimento, número de docentes e número de matriculados. Foram investigados 38 IF. Foram excluídos do estudo os dois CEFET, a UFTPR, o CPII e as ETF vinculadas às Universidades Federais. As bases utilizadas foram os relatórios anuais de gestão e os relatórios de governança e gestão de pessoas. Os dados que foram coletados e analisados compreendem apenas o ano de 2016. Eles concluíram sua análise destacando que: i) “existem poucos estudos que analisam as características da RFEPCT, sejam administrativas, educacionais, científicas, tecnológicas e extensionistas”; ii) foi possível agrupar os IF em três grupos homogêneos com base em suas características operacionais; e iii) a “dimensão” dos IF “não parece ser tão determinante em seus indicadores de eficiência”.

Ribeiro *et al.* (2020) utilizaram no seu estudo o Método Análise Não Hierárquica (em inglês, *Non Hierarchical Clustering* - NHC), especificamente o método de agrupamento *K-médias* (em inglês, *K-means*) para investigar a produção científica a partir da análise dos resumos baseado na frequência de palavras, para identificar pesquisas similares para agrupar os docentes de *campus* distintos do IFTM. A base de dados utilizada foram 370 resumos coletados do seminário de iniciação científica e inovação tecnológica do IFTM em 2017. Os resumos foram extraídos entre os meses de julho e agosto de 2018.

Menezes *et al.* (2019), ao utilizar uma análise estatística descritiva, avaliaram como a produção e Transferência Tecnológica ou de Tecnologia (TT), relacionada às patentes depositadas pelo IFES, ocorriam e propuseram estratégias para viabilizar sua transferência para sociedade. A base utilizada foi o INPI. As patentes foram coletadas entre 2012 e 2017. Ao final, eles concluíram que: i) a relação entre os investimentos e a realidade empírica suscita relevantes questionamentos; ii) os quantitativos de pesquisas realizadas não correspondem aos quantitativos de depósitos de patentes e iii) devem ser apresentadas estratégias para intensificar a proteção e a transferência de tecnologia, dentre as quais destacam-se o mapeamento de projetos de pesquisa com potencial inovador, a disponibilização de acesso ao banco de patentes e o estabelecimento de regras para publicação.

Miano *et al.* (2020) utilizaram-se de levantamento documental e análise estatística descritiva simples para investigar a produção científica desenvolvida pelos grupos de pesquisas atípicos cadastrados nas Escolas Técnicas do RJ, CEFET/RJ, IFF e IFRJ. A base de dados utilizada foi o DGP. Em suas conclusões, eles destacaram que: i) há uma carência de estudos que avaliem a trajetória quantitativa e qualitativa da produção científica dos grupos de pesquisa em atividade no ecossistema da RFEPCT do Rio de Janeiro; ii) há alta frequência de atipicidade dos grupos de pesquisa; e iii) há a necessidade de indicadores mais precisos envolvendo a produção científica dos grupos de pesquisa pertencentes à RFEPCT.

Após uma análise temporal e abrangente de estudos envolvendo os canais de PD&S da RFEPCT, ficou evidente que há a necessidade de identificar, organizar, monitorar e estabelecer diretrizes para esta Rede, assim como compreender suas características, passada sua primeira década. Contudo, essas análises são limitadas a uma ICT específica ou a um grupo restrito delas, tratando de casos isolados sem permitir a seleção de outros cenários a partir dos dados coletados. Além disso, a quantidade de dados

considerada nesses estudos é limitada quando comparada à abrangência do estudo proposto neste trabalho, que possibilita uma análise mais ampla e contextual de diversos cenários. Isso inclui a capacidade de comparar os resultados com aqueles obtidos anteriormente, a partir de várias dimensões e contextos.

A grande maioria dos estudos acima revisados, utiliza-se de análises basicamente envolvendo estudos bibliométricos e documentais, por exemplo Lino (2013); Niwa (2014) Bento (2015), Chaves et al. (2017), Miano et al. (2020). Esses estudos são geralmente analisados aplicando-se uma estatística descritiva simples. Esta estatística apresenta seus resultados com base no estudo de frequência, mas não inclui, por exemplo, análises estatísticas fundamentadas em testes de hipóteses, análises supervisionadas ou não supervisionadas, nem modelos de inferência ou preditivos. Essa abordagem busca compreender os contextos da Rede.

A revisão acima revelou uma extensa variedade de variáveis que representam a complexa diversidade dos dados associados aos canais de PD&S nas ICTs, particularmente na RFEPCT. Observou-se que as variáveis quantitativas mais frequentemente analisadas nesses estudos incluem financiamento, investimentos, número de patentes e volume de produção científica. Essas são obtidas para gestão, monitoramento e investigação, a partir de fontes de dados públicos como Plataforma Lattes, DGP, Relatórios Gerais das Instituições de Ensino e INPI.

A Plataforma Lattes, por exemplo, permite o acesso a informações detalhadas sobre pesquisadores e projetos de pesquisa, sendo amplamente utilizada em estudos como os de Garcia (2011), Perucchi *et al.* (2017) e Santos *et al.* (2015). O DGP, por sua vez, é uma fonte relevante para identificar grupos de pesquisa e suas áreas de atuação, como evidenciado Chaves *et al.* (2017), Garcia (2011), Macedo (2014) e Miano; Do Couto; *et al.* (2017 e 2020). Já o RGIE fornece informações sobre a gestão das ICTs, sendo útil para estudos que buscam analisar o impacto das políticas e práticas de governança, como os de Lino (2013) e Bento (2015). Por fim, o INPI é uma fonte importante para pesquisas relacionadas à propriedade industrial, como patentes e marcas, conforme demonstrado por Chaves *et al.* (2017) e Menezes *et al.* (2019).

O estudo procurou identificar quais canais foram comumente relatados, as respectivas variáveis consideradas, as soluções já desenvolvidas, as fontes de dados consultadas, os resultados já identificados e os possíveis *gaps* que podem servir como requisitos para o desenvolvimento do MVP (consultar apêndice, Quadro B. 5).

Contudo, é fundamental considerar variáveis menos examinadas, como a prestação de serviços, a desmotivação dos funcionários, a burocracia, a segurança jurídica e a lentidão nos procedimentos, sendo que alguns desses não foram escopo deste trabalho de investigação. Essas variáveis, possivelmente, podem ser fatores ocultos que impactam o desempenho das ICTs em termos dos canais de PD&S.

Por fim, a proposta deste trabalho de integrar e otimizar os dados relacionados aos canais de PD&S e de desenvolver uma solução tecnológica em BI voltada para apresentar esses dados é corroborada por Santos (2010), que já destacava a necessidade da RFEPCCT em investir no desenvolvimento ou aquisição de soluções tecnológicas especializadas que facilitem o acesso rápido e aberto do conhecimento produzido, além da necessidade de se avaliar com parcimônia a criação de novos periódicos científicos em formato eletrônico, pois a falta de clareza referente à gestão e ao monitoramento da produção científica e tecnológica e da prestação de serviço é um rico já constatado pelo TCU (2019a).

## **2.2 Desafios e Soluções para a Interação ICTs-SPP para o Desenvolvimento Socioeconômico**

No mercado competitivo do século XXI, o conhecimento é considerado a nova *commodity* (Hidalgo *et al.*, 2008). Neste cenário, espera-se que as ICTs sejam estratégicas na promoção da inovação e no desenvolvimento socioeconômico em que estejam inseridas (Etzkowitz *et al.*, 1995; Li *et al.*, 2021; Perkmann *et al.*, 2007).

Ao se analisar os dois últimos relatórios elaborados pela *Clarivate Analyticis*<sup>3</sup>, os quais abrangem o interstício de 2011 a 2018, constata-se que mais de 99% de toda produção científica brasileira é desenvolvida no âmbito das ICTs. Além disso, de acordo com o relatório, das 20 primeiras que mais publicam, cinco (5) são estaduais e quinze (15) federais, ou seja, são ICTs públicas. Em relação à produção tecnológica brasileira, sob a forma de patente (Peixoto *et al.*, 2021) contabilizada no INPI, as ICTs representam 28% do total.

Outra constatação que o referido relatório supracitado permitiu identificar foi a de que o Brasil até 2018 era o décimo terceiro no *ranking* internacional de produção de artigos científicos indexados (Analytics, 2019, p. 3), embora a partir de 2016 até 2018 tenha havido queda dessa produção. Por fim, ainda foi possível constatar, com base nos

---

<sup>3</sup> Os relatórios estão disponíveis nos seguintes sítios eletrônicos: <http://portal.andes.org.br/imprensa/noticias/imp-ult-992337666.pdf> e [https://jornal.usp.br/wp-content/uploads/2019/09/ClarivateReport\\_2013-2018.pdf](https://jornal.usp.br/wp-content/uploads/2019/09/ClarivateReport_2013-2018.pdf), ambos acessados em janeiro de 2019.

resultados das vinte (20) ICTs-MACE, entre 2011 até 2016, que a colaboração em publicações científicas envolvendo a interação ICTs-SPP no Brasil apresentou uma média 0,88% do total de publicações (Analytics, 2018, p. 42).

O relatório, também, constatou que embora nas últimas décadas a interação ICTs e SPP tenha sido uma importante política pública e de governo voltada ao desenvolvimento econômico e social no Brasil, relacionadas a interação TH, “muito poucas medições da quantidade e intensidade dessa interação foram realizadas” (Analytics, 2019, p.15).

Uma pesquisa realizada pelo Instituto de Pesquisa Econômica Aplicada IPEA (De Negri, F. O. *et al.*, 2016), em 2013, sobre as características do desenvolvimento tecnológico brasileiro, identificou que a infraestrutura de pesquisa brasileira necessita de melhorias e é limitada. Eles identificaram que a maior parte dos financiamentos para os canais de CT&I é proveniente do governo, por meio de instituições como o Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) e as Fundações de Apoio à Pesquisa (FAPs), além de grandes contribuições do SPP, como a Petrobras e a Financiadora de Estudos e Projetos (Finep).

Seus resultados sugerem também que no período investigado houve uma boa interação ICTs e SPP, com 43% das infraestruturas das ICTs prestando serviços às demandas dos SPP. Por fim, contataram que “a imensa maioria” da nossa infraestrutura de desenvolvimento tecnológico como laboratórios etc. estão localizados nas ICTs, em especial nas públicas. Portanto, esse estudo vem corroborar com os relatórios da *Clarivate Analytics* (Analytics, 2018, p. 42; 2019).

A elaboração das políticas direcionadas à PD&I é essencial para o fomento econômico e desenvolvimento regional de uma sociedade. Dentre as diversas políticas ocorridas entre 2008 e 2018, destaca-se o Plano de Ação 2007–2010 para Ciência, Tecnologia e Inovação (PACTI)<sup>4</sup>. Coincidentemente, ele foi proposto justamente no período em que a RFEPCT foi criada. Ao comparar os objetivos e finalidades da RFEPCT com as prioridades estratégicas propostas no PACTI, identifica-se uma convergência e alinhamento entre a política e sua execução.

Outras políticas foram propostas até 2018, dentre elas, por exemplo, a Lei 13.243/2016 (Brasil, 2016a), a qual dispõe sobre estímulos ao desenvolvimento científico à pesquisa, à capacitação científica e tecnológica e à inovação, denominada de marco

---

<sup>4</sup> [https://www.senado.gov.br/comissoes/cct/ap/AP20080417\\_MCTMinSergioRezende\\_PlanoAcao.pdf](https://www.senado.gov.br/comissoes/cct/ap/AP20080417_MCTMinSergioRezende_PlanoAcao.pdf), acesso em 21 mar 2017.

legal da inovação. De Carvalho *et al.* (2020) já identificavam em sua pesquisa indícios de uma cultura aversa à inovação no Brasil, justamente pelos três ecossistemas que compõem a TH. Recentemente, Rossi (2022) constatou que a falta de recursos humanos especializados e competentes na transferência de tecnologia gera insegurança e demora no processo dessa transferência dos produtos, processos e serviços decorrentes dessa interação.

Em 2022, o Tribunal de Contas da União (TCU) emitiu uma lista de alto risco da administração pública sobre vários temas, dentre eles o da política de inovação e a qualidade, compartilhamento, padronização e transparência dos dados governamentais (Brasil, 2022a). Essa lista identificou falhas em todo o ciclo das políticas de CT&I que podem comprometer a transformação dos avanços científicos em ganhos de produtividade e competitividade para o país.

A padronização de dados (Tong *et al.*, 2021) refere-se à uniformização dos procedimentos de coleta, classificação, armazenamento e análise de dados, de modo a garantir que as informações sejam comparáveis e confiáveis. Isso implica definir previamente as variáveis de interesse, as unidades de medida, os formatos e os padrões de apresentação dos dados, visando minimizar possíveis erros de interpretação e inconsistências nas análises.

A transparência de dados (Giroto *et al.*, 2020) refere-se ao grau de acessibilidade e clareza das informações disponibilizadas ao público em geral. Um sistema transparente de dados deve permitir que as informações estejam disponíveis de forma acessível, clara e objetiva, sem distorções ou ocultamentos que possam prejudicar a tomada de decisões informadas e a participação democrática da sociedade.

A falta de padronização (Pereira, 2022) e transparência nos dados públicos (Brasil, 2022a, p. 108-111) pode explicar a diferença entre as variáveis consideradas neste trabalho e as identificadas na literatura, por exemplo quando se trata de patente está se referindo a patente depositada, publicada, concedida ou transferida? Assim, embora os conceitos de padronização e transparência de dados estejam relacionados, eles se referem a aspectos distintos na gestão e na divulgação de informações públicas. A padronização é fundamental para garantir a confiabilidade dos dados, enquanto a transparência é fundamental para garantir o acesso e a compreensão desses dados pela sociedade. Importa salientar, no entanto, que existe um padrão inconsistente sobre o que essas variáveis efetivamente medem, incongruência esta observada inclusive na maioria dos estudos analisados.



Na literatura consultada, por exemplo, foi possível identificar que, quando a variável "*projetos de pesquisa*" é utilizada, não fica claro se ela considerou apenas projetos ativos, concluídos, encerrados ou qualquer combinação dessas possibilidades ou, até mesmo, sua diferenciação entre básica ou aplicada (cuja maioria da sua classificação pelos gestores ou pesquisadores dá-se por critérios subjetivos). Portanto, a diferença encontrada entre as variáveis identificadas na literatura e as variáveis consideradas no trabalho pode ser explicada pela deficiência na padronização e transparência dos dados disponibilizados.

Nesse sentido, é crucial realizar pesquisas e desenvolver soluções que estabeleçam critérios claros e precisos para a definição dessas variáveis, permitindo uma análise mais consistente e confiável dos dados (Brasil, 2022a, p. 50-53; 108-111).

As deficiências relacionadas ao ecossistema da CT&I e PD&I (Brasil, 2022a, p. 50-53) incluem a falta de diagnóstico adequado dos problemas, o excesso de temas prioritários, a ausência de coordenação efetiva entre os atores ICTs, SPP e GOV, a falta de metas claras e de indicadores de resultado, a ausência de planejamento estratégico de longo prazo e as suas fragmentações, sobreposições e duplicações de projetos científicos e ou tecnológicos.

As deficiências relacionadas à qualidade, compartilhamento e transparência dos dados públicos (Brasil, 2022a, p. 108-111) incluem a inconsistência e insuficiência dos dados avaliados, a falta de validação dos critérios das políticas públicas devido à ausência de consultas a bases de dados de outros órgãos, a formulação, avaliação e monitoramento de políticas incompletas ou deficientes devido à falta de acesso a bases de dados de outros órgãos, a baixa integração entre sistemas de informação, a dificuldade de acesso às bases de dados de outros órgãos, a insuficiência de controles de sistema de gestão e de rotina de tratamento dos dados e um catálogo de bases de dados incompleto.

Corroborando com o TCU, Pereira (2022) concluiu ser comum que dados públicos apresentem problemas de estrutura, padronização, dificuldades de localização, burocracia no acesso, incompletude de registros, colunas ausentes ou, em alguns casos, sendo necessário utilizar a lei de acesso à informação (Brasil, 2011) quando os dados de um determinado ano não estão disponíveis.

O contexto apresentado indica a necessidade de soluções para mitigar problemas relacionados à falta de diagnóstico adequado e insuficiência de dados avaliados neste ecossistema. Para isso, sugere-se a utilização de soluções de BI, que podem ser aplicadas para análise e interpretação dos dados coletados. Com o uso dessa ferramenta, será

possível avançar em outros estudos, utilizando técnicas de estatísticas multivariadas supervisionadas e não supervisionadas para uma análise mais aprofundada dos dados.

Portanto, para que haja interação nos canais de PD&S e possivelmente inovação, é necessário não apenas investimento, público ou privado, mas também ações colaborativas de múltiplos atores. No entanto, (Niwa, 2014) afirmou que a burocracia, lentidão nos procedimentos e desmotivação dos docentes e técnicos são fatores que podem impedir essa colaboração eficiente.

Urge-se necessário e imediato apresentar soluções que possam contribuir para a resolução ao menos de parte dos problemas listados, os quais impactam diretamente em ecossistema de CT&I e PD&I do Brasil.

### **2.3 Revisão da Evolução da Teoria Tripla Hélice de Inovação**

Conceitualmente, a TH apresenta como tese (Carayannis *et al.*, 2022; Etzkowitz *et al.*, 2000; Galvão *et al.*, 2019) que a universidade pode desempenhar um papel mais importante, com potencial de estimular a inovação e o desenvolvimento regional em sociedades nas quais o conhecimento ocupa um papel mais proeminente.

Essa proeminência deve ocorrer sem que haja sobreposição, nem destaque, no arranjo institucional científico e tecnológico (universidades, institutos, centros e laboratórios de pesquisas e desenvolvimento tecnológico), setor produtivo e governo (Carayannis *et al.*, 2022).

Doravante, essa sobreposição foi sintetizada como Universidade-Industria-Governo (UIG) com o objetivo de gerar novos formatos institucionais e sociais que fomentassem a produção, transferência e aplicação do conhecimento científico e tecnológico resultando, em alguns casos, em inovação.

A Quadrupla Hélice (Carayannis *et al.*, 2009; Park, 2014) contextualiza a TH, adicionando como a quarta hélice “sociedade civil” e o “público baseado na mídia e na cultura” e a Quintupla Hélice (Carayannis *et al.*, 2010; Park, 2014) incorpora a Quadrupla Hélice, estendendo a arquitetura de inovação ao meio ambiente e à ecologia social. Embora, o crescimento de proposições de modelos posteriores a TH decorram da complexidade com que essa interação evolui socialmente, legalmente e tecnologicamente, e, contrário a essas propostas, Leydesdorff (2012) destacou que “*so long as one is not able to operationalize and show development in the relatively simple*

*case of three dimensions, one should be cautious in generalizing beyond the TH model to an N-tuple of helices*”;<sup>5</sup> para mais, consultar (Leydesdorff, 2013).

A TH é um modelo analítico que busca explicar a dinâmica dos sistemas de inovação em seus diversos níveis (local, regional, nacional e internacional), considerando a variedade de arranjos institucionais e modelos de políticas com o objetivo de desenvolver um ambiente empresarial inovador, o conhecimento como gerador econômico e arranjos complexos envolvendo, no mínimo, a interação entre a UIG (Etzkowitz *et al.*, 2000; Leydesdorff *et al.*, 1996; 1998). Esse modelo é sugerido quando o objetivo visa mensurar a maturidade inovativa de modo sistêmico, assim não é necessário assumir a existência prévia de sistemas nacionais (ou regionais) de inovação.

Há uma outra linha teórica que analisa apenas a colaboração universidade-indústria, em inglês UIC. A UIC foi amplamente percebida como uma ferramenta promissora para aumentar a capacidade organizacional em inovação aberta (Ankrah *et al.*, 2015) e apresenta como campo de estudo à interação entre qualquer parte do sistema educacional superior e a indústria visando principalmente incentivar intercâmbio de conhecimento e tecnologia (Ankrah *et al.*, 2015; Bekkers *et al.*, 2008; Siegel *et al.*, 2003).

Embora possa haver um aumento da pesquisa relacionada à UIC, a literatura existente ainda é relativamente fragmentada e não apresenta uma visão abrangente (Ankrah *et al.*, 2015; Bovaird, 2007; Perkmann *et al.*, 2013). Enquanto a TH é um modelo analítico, a UIC é uma ferramenta; enquanto a TH busca explicar a dinâmica dos sistemas de inovação e desenvolvimento econômico em seus diversos níveis, a UIC busca aumentar a capacidade organizacional em inovação aberta; e, enquanto na TH há interação, a UIG visa gerar novos formatos institucionais e sociais para a produção, transferência e aplicação do conhecimento. Por fim, na UIC a interação entre qualquer parte do sistema educacional superior e a indústria visa principalmente incentivar intercâmbio de conhecimento e tecnologia.

No cenário, o *locus* de análise que envolve a relação entre UIG possui um modelo e uma ferramenta que nos ajuda a compreender a evolução dessa complexa rede. Ao se consultar as duas principais bases de indexação *Scopus* e *Web of Science* percebe-se que, historicamente, a interação entre UIG vem apresentando um crescimento constante e ascendente, desde o ano 2000. Essa afirmação pode ser constatada ao verificar o resultado

---

<sup>5</sup> Numa tradução livre: “contanto que não seja capaz de operacionalizar e mostrar o desenvolvimento no caso relativamente simples de três dimensões, deve-se ser cauteloso ao generalizar além do modelo TH para uma *n-upla* de hélices”.

do gráfico de linha gerado por essas bases, supracitadas, ao buscar pela palavra-chave *UIG, University-Industry-Government* e suas variações. Esse resultado pode ser atestado ao contabilizar a quantidade de documentos publicados que citaram o artigo semente (Etzkowitz *et al.*, 2000). O artigo semente representa um avanço científico ou uma descoberta transformadora. Neste caso, compreende-se que houve um avanço científico, ao propor o modelo TH e uma mudança transformadora em como compreender a interação universidade, governo e indústria, no qual todos passam a ter a mesma importância.

A evolução sintetizada surgiu da necessidade de, enquanto campo investigativo, identificar e revisar os modelos conceituais anteriores (Apêndice B, Quadro B. 4)

A teoria da TH destaca a importância das interações entre ICTs, SPP e GOV para o desenvolvimento de inovação e crescimento econômico (Etzkowitz *et al.*, 2000). Isso está em consonância com alguns dos objetivos e finalidades da RFEPCT (Brasil, 2008), que também busca promover a inovação e o desenvolvimento socioeconômico.

Neste cenário complexo e diversificado do modelo TH, este trabalho buscou-se, não apenas propor a criação de uma ferramenta, mas também, uma análise da maturidade relacionada aos canais PD&S da RFEPCT. Tradicionalmente, os estudos empíricos abordando as análises dos canais CT&I e PD&I, considerando o modelo TH como abordagem analítica, podem ser organizadas em duas áreas distintas, porém complementares.

A primeira, fundamenta-se no conceito da entropia da informação, proposto por Shannon (1948). Dentre os estudos aplicando esse conceito para estudar as interações TH pode destacar, por exemplo Leydesdorff *et al.* (2006) e Barragan-Ocana *et al.* (2020).

A segunda apoia-se no conceito da informação multivariada proposto por McGill (1954). Dentre os estudos aplicando esse conceito para estudar as interações TH pode destacar, por exemplo Kiyani *et al.* (2018), Sinha *et al.* (2019) e Yaakub *et al.* (2020).

Sobre a primeira, contata-se que os estudos buscam investigar a sinergia entre os atores em seus diferentes níveis e contextos. Por exemplo, Leydesdorff *et al.* (2006) sistematizaram a primeira proposta para medir a economia do conhecimento de uma determinada região, utilizando a TH. Recentemente, Barragan-Ocana *et al.* (2020) propuseram um índice alternativo de inovação tecnológica, agrupando indicadores reportados pela Organização para a Cooperação e Desenvolvimento Econômico (OCDE) como subíndices de entradas e saídas para classificar um grupo de países.

O foco de investigação sobre a RFEPECT, neste estudo, adotou a segunda abordagem. Essa abordagem investiga os resultados, interações, performances e características envolvendo um ou mais atores da TH.

Vários estudos empíricos, adotando a segunda abordagem, foram conduzidos com o objetivo de investigar diferentes aspectos da interação ICTs-SPP. Kiyani *et al.* (2018) realizaram *exploratory factor analyses* para validar o comportamento de cidadania organizacional entre acadêmicos de ICTs privadas no Paquistão. Sinha *et al.* (2019) abordaram a satisfação do SPP ao empregar recém-formados em engenharia na Índia. Já Adeinat *et al.* (2019) examinaram o impacto da cultura de uma ICTs na GC. Enquanto Sierra *et al.* (2018) investigaram as variáveis relevantes na TT e interação ICTs-SPP de uma ICTs pública do México. Adicionalmente, Yaakub *et al.* (2020) mediram a performance de uma ICTs particular na Malásia, em relação ao ranqueamento acadêmico.

Jackson *et al.* (2018) utilizaram *cluster analysis* para identificar características de colaboração em inovação comparando resultados da Austrália com os da OCDE. D'este *et al.* (2007) empregaram *exploratory factor analyses* para examinar fatores que influenciam o envolvimento dos pesquisadores em interações ICTs-SPP. Cesaroni *et al.* (2016) realizaram *cluster analysis* e *multinomial logistic regression* para analisar ICTs italianas, investigando migração de modelos de transferência de conhecimento e fatores explicativos de diferenças entre as ICTs analisadas. Berbegal-Mirabent *et al.* (2019) categorizaram ICTs espanholas através de *cluster analysis*, considerando atividades de ensino, pesquisa e TT. Adicionalmente, Fan *et al.* (2015) agruparam ICTs chinesas através de *cluster analysis* baseada em colaboração ICTs-SPP voltada à pesquisa, coautoria de acadêmicos e pesquisadores da indústria em periódicos internacionais. Além disso, Jackson *et al.* (2018) utilizaram *cluster analysis* para analisar (de)similaridades entre Austrália e Reino Unido, decorrentes de colaboração ICTs-SPP.

Nas últimas duas décadas, as ICTs iniciaram a atualização de suas práticas, influenciadas pelas demandas da TH. Compreender os canais de PD&S ao longo do desenvolvimento institucional é essencial para atender às demandas setoriais com qualidade profissional e social adequada (Cabelkova *et al.*, 2017; Cesaroni *et al.*, 2016; Ranga *et al.*, 2013).

Apesar de muitos estudos já terem sido conduzidos, são poucos os que visam compreender a consolidação das atividades das ICTs. Essa observação é corroborada por Kim *et al.* (2009), que identificaram a necessidade de analisar o desempenho das ICTs decorrente de seus canais. Čábelková *et al.* (2017) propuseram examinar o papel das ICTs

no desenvolvimento de *clusters* industriais em regiões periféricas, destacando a necessidade de mais estudos empíricos para explicar as diferenças no desenvolvimento regional global. Além disso, Ramos-Vielba *et al.* (2010) constataram que estudos analisando essa interação apresentam limitações importantes devido à escassez e inconsistência dos dados empíricos disponíveis. Destaca-se que estudos similares envolvendo ICTs ou o SNI no Brasil, são limitados tendendo a escassos.

Vários estudos anteriores propuseram a criação de indicadores agregados em diferentes contextos. Yaakub *et al.* (2020) constataram que muitas ICTs buscam medidas adequadas de capacidade de inovação, identificando três principais variáveis: despesas governamentais de P&D, desenvolvimento tecnológico e fatores internos e externos que afetam os sistemas de inovação. Villasenor *et al.* (2017) identificaram que o monitoramento da pesquisa acadêmica, incluindo classificações e indicadores cientiométricos, é um tópico essencial nas agendas científicas, embora ferramentas computacionais mais sofisticadas raramente sejam exploradas. Tijssen *et al.* (2009) destacam a deficiência de informações comparativas entre ICTs baseadas em seus canais de pesquisa e prestação de serviços para os SPP, alertando que as análises estatísticas utilizadas para estabelecer indicadores devem ser preferencialmente aplicadas em estudos de *benchmarking* multidimensionais não avaliativos, não para ranqueamento universitário. D'este *et al.* (2007) destacam a negligência em relação a questões importantes, como projetos conjuntos de pesquisa, consultoria e treinamento, apesar da variedade de canais relacionados à transferência do conhecimento.

A revisão da literatura (Seção 2) possibilitou o mapeamento de 43 variáveis/métricas (Apêndice B, Quadro B. 5), a compreensão da interação entre os atores (Apêndice B, Quadro B. 2) e a efetividade da colaboração (Seção 4).

Avaliar o desempenho das políticas e ações, relacionadas aos canais de PD&S, permite identificar características favoráveis, desafios e oportunidades nos ecossistemas de inovação.

Essa revisão de literatura contribuiu para a compreensão da TH no contexto brasileiro de uma rede de ICTs, à medida que pode ser utilizada para subsidiar decisões estratégicas e aprimorar as ações dos atores envolvidos para potencializar os resultados e impactos gerados na inovação e pesquisa, fortalecendo assim a posição do Brasil no cenário global de tecnologia e ciência. Além disso, ao alinhar as práticas e decisões com *insights* derivados dessa análise, é possível criar um ambiente mais colaborativo e

eficiente, onde os atores da rede de ICTs trabalham juntos para superar desafios e aproveitar oportunidades emergentes.

#### **2.4 Revisão da Literatura sobre *Business Intelligence*, *Data Warehouse*, *Dashboards* e Indicadores-chave de Desempenho Aplicados ao Ecossistema de ICTs**

A utilização de técnicas de *Business Intelligence* (BI) e de tecnologias avançadas de gestão de dados tem se destacado como uma ferramenta essencial no panorama atual da informação. Estas soluções, quando aplicadas de forma estratégica, oferecem um potencial inovador, otimizando a forma como os dados são visualizados, analisados e interpretados, especialmente em campos complexos de atuação, como a TH que abrange a interação entre ICTs, SPP e GOV. Nesse contexto, a junção da técnica de BI a tecnologias como *dashboard* e *Data Warehouse* (DW) emerge, não apenas como uma resposta às demandas de gestores ou órgãos de controle governamentais, mas também como um recurso vital para pesquisadores e a sociedade como um todo, garantindo que a informação seja precisa, relevante e de fácil acesso.

Segundo o sociólogo e economista austríaco Peter Ferdinand Drucker (1909-2005), a sociedade do conhecimento (Lipkowitz, 1969) transformou a maneira como as organizações operam, particularmente no domínio das ICTs. Drucker postulou que o conhecimento e a informação substituem consideravelmente recursos físicos e mão-de-obra como motores principais do desenvolvimento socioeconômico, sendo cruciais para o sucesso na nova realidade econômica.

Ferramentas tecnológicas como *dashboards* e *Data Warehouse*, em conjunto com técnicas de BI (Devlin, 1996), têm encontrado amplo uso em diversas áreas do conhecimento, como ecologia e administração (da Cunha *et al.*, 2023; Kongthanasuwan *et al.*, 2023). Essas ferramentas são fundamentais para a análise integrada de informações, favorecendo um entendimento mais amplo e eficaz dos dados relevantes. Este conjunto de recursos promove uma gestão mais efetiva de informações e proporciona *insights* valiosos para tomada de decisões em diversos setores.

O atual cenário caracteriza-se pelo uso significativo dessas tecnologias, contribuindo especialmente para áreas como ecologia e administração, ao oferecer *insights* para orientar decisões estratégicas. É evidente o seu potencial em proporcionar benefícios para diferentes áreas de estudo. Dentre as ICTs, a gestão e análise de grandes

volumes de dados internos e externos auxilia na tomada de decisões táticas, estratégicas e operacionais, embora essa transformação de dados brutos em informações demande por tecnologias sofisticadas como *dashboards* e *Data Warehouse*.

Atualmente em associação com tais tecnologias, os sistemas baseados em BI têm emergido como uma abordagem promissora para apoiar a tomada de decisões baseada em dados nas ICTs (Azevedo *et al.*, 2021; Damyanov *et al.*, 2019; Hoffmann, 2021; Medina *et al.*, 2018; Santos, Thiago Rizzi *et al.*, 2021; Zhai *et al.*, 2010).

A técnica de BI emerge como uma abordagem promissora para apoiar essa tomada de decisões baseada em dados, sendo crucial para a eficiência operacional, a inovação e a qualidade do ensino e da pesquisa nas ICTs. Entretanto, apesar do seu potencial considerável, a compreensão de sua adoção e implementação nas ICTs, especialmente quando aplicada aos canais de PD&S, ainda é limitada (Jamiu *et al.*, 2020; Santi *et al.*, 2018; Santos, T. R. *et al.*, 2021), seja em relação à qualidade dos dados inseridos nas fontes de dados, à falta de sistematização de dados, à falta de modelos de sistemas ou a canais ainda a serem investigados.

Este estudo objetiva analisar a adoção e aplicação da técnica de BI no ecossistema das ICTs. Por meio de revisão literária e análise de estudos de caso relevantes, busca-se identificar os fatores que influenciam a adoção dessa técnica, as estratégias para superar desafios associados e seu impacto na eficiência operacional e inovação. Visa também fornecer previsões valiosas para pesquisadores e profissionais interessados no potencial dos sistemas web baseados em BI para transformar a gestão e operação das ICTs, além de contribuir para a literatura existente sobre a adoção de Tecnologias de Informação e Comunicação no setor de educação e pesquisa.

#### **2.4.1 Histórico e Evolução de BI**

Conforme Srivastava *et al.* (2022), a técnica de BI emergiu e sofreu transformações significativas desde sua concepção pelo editor americano Richard Miller Devens (1824 - 1900), em 1865, para atender às necessidades do mercado e tecnologias. Duas ondas principais de aplicações surgiram: a primeira centrada no processamento interno das informações das organizações e a segunda no desenvolvimento de software baseado na *web* e *Big Data* (Santos, T. R. *et al.*, 2021).

Heang (2017) destacou a evolução paralela da tecnologia de suporte à decisão em conjunto com a técnica de BI, que se tornou fundamental para as decisões estratégicas,



operacionais e táticas, especialmente com o aumento da coleta de dados e avanços na tecnologia de armazenamento.

A década de 1990 viu inovações como o *Data Warehouse* (DW) e tecnologias de Processamento Analítico Online (em inglês, *Online Analytical Processing* – OLAP), facilitando o armazenamento eficiente e análises complexas (Srivastava *et al.*, 2022).

O início do século XXI marcou a emergência da técnica moderna de BI, com implementações avançadas, incluindo a *data mining* e os painéis (em inglês, *Dashboards*), com Inteligência Artificial (em inglês, *Artificial Intelligence* – AI) desempenhando um papel propulsor (Ain *et al.*, 2019) no desenvolvimento de sistemas, plataformas e observatórios baseados em BI, particularmente na área de análise preditiva e automação de processos.

As recentes inovações permitiram às organizações explorar seus dados para descobrir padrões anteriormente ocultos e obter *insights*, conduzindo a mudanças significativas nos processos decisórios (Ahmed Muayad *et al.*, 2022; Santos, T. R. *et al.*, 2021). Essa exploração potencializou a eficiência organizacional.

As técnicas de DM, fundamentais nessa descoberta de padrões, contribuíram para facilitar uma melhoria no desempenho organizacional (Santos, T. R. *et al.*, 2021), integrando-se à exploração dos dados, de maneira a aprimorar a análise e a interpretação.

O progresso de BI foi reforçado por teorias como a da informação, decisão e gestão do conhecimento, facilitando a maximização do valor dos dados e gerando vantagens competitivas (Ain *et al.*, 2019; Heang, 2017; Musa *et al.*, 2019; Santos, T. R. *et al.*, 2021; Srivastava *et al.*, 2022; Tripathi *et al.*, 2020). Estas teorias contribuem para o aprimoramento do processo decisório estratégico (Ain *et al.*, 2019).

No campo educacional, a importância dos sistemas de BI tem sido evidenciada, especialmente em canais administrativos, financeiros e de ensino (Santi *et al.*, 2018). Esses sistemas desempenham um papel crucial às ICTs, permitindo transformar seus dados brutos em informações valiosas. Tal transformação auxilia na orientação de operações, inovações e estratégias, fomentando assim a tomada de decisões eficazes (Santos, T. R. *et al.*, 2021). Exemplos específicos da adoção de técnicas de DM e a aplicação de sistemas de BI, nas ICTs, consultar Apêndice B, Quadro B. 1.

A governança de dados torna-se componente essencial, garantindo a qualidade e confiabilidade das informações nos sistemas baseados em BI, sendo primordial para a tomada de decisões que podem impulsionar o sucesso das ICTs e o desenvolvimento socioeconômico. Ou seja, é vital não negligenciar a qualidade dos dados manipulados.

## 2.4.2 Teorias Relevantes em BI

No campo de BI, teorias relevantes fundamentam a compreensão, aplicação e aprimoramento da disciplina, permitindo a análise de dados e a otimização do desempenho organizacional. A adequada combinação e aplicação dessas e outras teorias têm o potencial de conferir às organizações uma vantagem competitiva, transformando dados brutos em conhecimento aplicável e embasando em dados possibilitando decisões estratégicas mais assertivas.

Várias perspectivas acadêmicas influenciam a BI. A seguir são apresentadas as principais teorias que norteiam o desenvolvimento de soluções baseadas em BI.

### 2.4.2.1 Teoria da Decisão Racional

A Teoria da Decisão Racional (TDR), de acordo com Heukelom (2010), descreve o processo pelo qual indivíduos ou organizações tomam decisões ótimas ao avaliar várias alternativas. Essa teoria, centrada em axiomas lógicos, estabelece regras para deliberar sobre proposições, guiando tanto indivíduos quanto organizações a optar por decisões ótimas, após avaliação de diversas alternativas. Essencial no domínio dos negócios, a TDR tem papel destacado no desenvolvimento de sistemas de BI com o intuito de respaldar decisões estratégicas (Mls *et al.*, 2009). Como enfatizado por Mls *et al.* (2009), ao possuírem informações completas, os decisores tendem a favorecer a opção que maximiza sua utilidade esperada.

### 2.4.2.2 Teoria da Gestão do Conhecimento

A Teoria da Gestão do Conhecimento (TGC), segundo Sumner (1999), foca na conversão do conhecimento tácito, intrinsecamente pessoal e subjetivo, em uma forma explícita e compartilhável. Esse processo facilita a capacidade das empresas de materializar ideias em inovações tecnológicas e produtos. Paralelamente, Tzortzaki *et al.* (2014) destacou a relevância de BI como ferramenta crucial na coleta e análise de dados, com o intuito de fundamentar decisões. Ao integrar as capacidades da TGC e de BI, objetiva-se potencializar o conhecimento, otimizando assim o processo decisório.

#### 2.4.2.3 Teoria da Análise de Negócios

A Teoria da Análise de Negócios (TAN) propõe um conjunto de estratégias e ferramentas para identificar e analisar necessidades e requisitos empresariais, buscando aprimorar a eficiência de processos (Chasin, 2016). Tal abordagem auxilia na compreensão de fatores cruciais para o êxito em setores como o desenvolvimento de novos produtos e operações empresariais (Kashefi *et al.*, 2018; Liu, 2020). No contexto de BI, aplicar esta teoria implica no uso de técnicas como *data mining*, análise preditiva e modelagem estatística, com o objetivo de extrair *insights* valiosos (Chasin, 2016; Kashefi *et al.*, 2018; Perera *et al.*, 2012). Estes *insights* possibilitam às instituições discernir tendências, padrões e oportunidades de mercado, culminando em decisões estratégicas mais embasadas.

#### 2.4.2.4 Teoria da Inteligência Competitiva

A Teoria da Inteligência Competitiva (TIC), descrita por Garcia-Alsina *et al.* (2012) e Breakspear (2013), é uma abordagem multidisciplinar que busca coletar, analisar e disseminar informações do ambiente competitivo. Aliada ao BI, essa abordagem aperfeiçoa a tomada de decisões estratégicas, ao monitorar concorrentes e tendências setoriais (Negro *et al.*, 2020; Wang, 2023). O processo respeita princípios éticos e legais, focando em dados publicamente disponíveis e extraídos por métodos invasivos (Breakspear, 2013). A interdisciplinaridade fortalece sua aplicação e é essencial para a competitividade em mercados complexos (Garcia-Alsina *et al.*, 2012).

#### 2.4.2.5 Teoria da Visualização de Dados

A Teoria da Visualização de Dados (TVD), respaldada por Palloni *et al.* (2018, p. 32-38) e Gavrilova *et al.* (2017), visa transformar dados brutos em representações gráficas claras, utilizando-se de *design* gráfico, psicologia perceptual e ciência cognitiva. Esse campo, ao se integrar a BI, tem o propósito de facilitar a identificação de padrões e *insights* em grandes volumes de informação. Tais representações gráficas se mostram essenciais para permitir decisões bem fundamentadas no âmbito da inteligência de negócios contemporânea.

Considerando que: as TAN e da TVD são amplamente reconhecidas na literatura científica, por seu valor estratégico e aplicabilidade na gestão de informações e decisões empresariais; que ambas apresentam uma abordagem integrativa, com a TAN fornecendo ferramentas e estratégias para a compreensão profunda das necessidades e requisitos organizacionais, e a TVD transformando dados brutos em representações gráficas claras; que as soluções em BI demandam uma assimilação rápida e eficaz das informações, minimizando ambiguidades e maximizando a clareza na tomada de decisão; e que os objetivos norteadores da gestão de dados dos canais de PD&S em ICTs estão em consonância com os princípios das teorias mencionadas, ambas serão utilizadas neste trabalho como fundamentos teóricos em associação com a teoria da TH. Essas teorias proporcionarão uma compreensão abrangente de como a solução tecnológica proposta em BI, ao ser aplicada ao caso de ICTs, pode integrar, otimizar e aprimorar a tomada de decisões, atendendo assim às exigências e demandas multifacetadas dos *stakeholders* envolvidos.

### **2.4.3 Definição e Componentes de BI**

A técnica de BI, constitui-se como uma abordagem analítica integrada, englobando processos, metodologias e tecnologias destinadas a coletar, analisar e apresentar dados de várias fontes. Seu propósito reside em criar *insights* acionáveis para a tomada de decisões estratégicas, táticas e operacionais, com o objetivo de aprimorar a eficiência operacional e identificar novas oportunidades de negócio, conferindo uma vantagem competitiva (Balakrishnan, 2022; Devlin, 1996).

Dentre seus componentes, destacam-se o Banco de Dados, os DW, o OLAP, a DM, as Consultas, Relatórios, os *Dashboards* e os *Scorecards*. A técnica de BI visa identificar as fontes de dados (em inglês, *datasources*), para, em seguida, extraí-las, transformá-las, carregá-las (ETL) e transferi-las a um repositório central de informações DW (Fana *et al.*, 2021; Krmac, 2011). Após a realização dessas etapas, os resultados possíveis incluem o processamento OLAP, a DM, o *Dashboards*, entre outros (Subseção 3.5).

#### 2.4.3.1 Bancos de Dados

Os Bancos de Dados (BD) constituem a base essencial para o desenvolvimento de soluções web em BI, sendo responsáveis por armazenar os dados operacionais e históricos da organização. Os BD podem incluir informações sobre diversas métricas, como artigos, produção tecnológica, patentes e pesquisas, aspectos considerados relevantes. Otimizados para consultas eficientes e para o gerenciamento de grandes volumes de dados, esses BD são comumente representados e estruturados pelo Modelo de Entidade Relacional (em inglês, *Relational Entity Model* – REM) (Hampel, 2021).

O REM é empregado para representar a estrutura dos dados, o que possibilita a organização lógica e visualização das entidades e suas interconexões no Sistema de Informação Gerencial (SIG). Utilizado na fase de projeto de banco de dados, o REM auxilia na criação da estrutura lógica que será implementada em um Sistema Gerenciador de Banco de Dados (SGDB) e manipulada por meio da Linguagem de Consulta Estruturada (em inglês, *Structured Query Language* – SQL) (Hampel, 2021).

#### 2.4.3.2 Extração, Transformação e Carga

O processo de Extração, Transformação e Carga (ETL) é a base para a construção de uma solução de BI eficiente. Ele compreende três etapas distintas: extração, transformação e carga. Na fase de extração, dados são coletados de diversas fontes, como bancos de dados, planilhas e sistemas externos. Em seguida, na etapa de transformação, os dados são limpos, padronizados e integrados para garantir a consistência e a qualidade das informações. Por fim, na etapa de carga, os dados são carregados em um conjunto de repositório de dados simples (em inglês, *Data Mart*) e, na sequência, salvos em uma *Data Warehouse*, prontos para a análise.

O processo de análise de dados, foi realizado o procedimento de ETL. O objetivo principal do ETL é integrar fontes heterogêneas de dados em um DW, conforme descrito por autores como Kimball *et al.* (2008), Diouf *et al.* (2018) e Ammar *et al.* (2022). Neste trabalho, os procedimentos de ETL foram adaptados a partir da metodologia proposta por Paikaray *et al.* (2021, p. 505-508).

O processo de ETL é uma etapa crucial no fluxo de trabalho analítico em projetos de BI. Ele constitui entre 60% a 80% do tempo dedicado ao projeto (Eckerson *et al.*,

2003). Durante essa etapa, são realizadas inúmeras transformações e alterações para preparar os dados antes de carregá-los no DW (Paikaray *et al.*, 2021).

Dentre os procedimentos disponíveis, destacam-se a limpeza, duplicação e revisão de formato, entre outros (Fikri *et al.*, 2019; Paikaray *et al.*, 2021).

O procedimento técnico denominado de *web scraping*, extração ou raspagem dos dados, consiste em ler e filtrar os dados disponíveis nas diferentes fontes de dados para uso em projetos de BI (Fu, 2016), dentre outros. Os dados raspados podem ser estruturados e não estruturados (Santana, 2019). Estruturados são aqueles que contém uma organização para serem recuperados (por exemplo: *xls, csv, etc*), ou seja, são organizados em linhas e colunas. Não estruturados são aqueles que apresentam uma organização não explícita de disposição (por exemplo: *pdf, html, etc*), ou seja, não são organizados em linhas e colunas. A raspagem para este trabalho ocorreu de maneira automatizada.

#### 2.4.3.3 Data Warehouse

Um *Data Warehouse* é um repositório centralizado e integrado que armazena dados de diversas fontes, incluindo sistemas transacionais, bancos de dados operacionais e até mesmo dados externos. Essa estrutura permite que os dados sejam organizados de forma coerente e histórica, possibilitando análises mais complexas e amplas em comparação com bancos de dados transacionais. Ammar *et al.* (2022) definiram DW como um banco de dados multidimensional otimizado para consultas analíticas de diferentes fontes independente do conjunto ou coleção utilizada para implementar uma solução em BI. Ou seja, um DW possibilita organizar uma grande quantidade de informações institucionais para tomada de decisões, prospecção e desenvolvimento de oportunidade de modo dinâmico e integrado agregando-lhe operações analíticas (Ammar *et al.*, 2022; Fana *et al.*, 2021).

A elaboração de um DW (Barahama *et al.*, 2021; Fana *et al.*, 2021; Singh *et al.*, 2021), ou modelo multidimensional, depende da escolha de como organizar uma ou mais Tabelas Dimensões e Fatos, aplicando-lhes uma modelo estrela (em inglês, *star schema*) e ou floco de neve (em inglês, *snowflake schema*).

O DW é uma tecnologia que tem sido amplamente utilizada em organizações de diversos setores para gerenciamento de grandes volumes de dados. Ela consiste na organização dos dados em esquemas (Hassan *et al.*, 2021) que permitem uma melhor

visualização e análise dos dados. Nesse contexto, destaca-se e os esquemas mais comuns são denominados estrela e floco de neve (Fana *et al.*, 2021; Iqbal *et al.*, 2020), que apresentam características distintas e que devem ser avaliadas de acordo com as necessidades específicas de cada projeto.

O esquema estrela, por exemplo, é composto por um pequeno número de tabelas dimensão e fato, cujos relacionamentos suportam consultas mais rápidas e com maior consistência de dados, conforme destacado por Japal *et al.* (2021).

O esquema floco de neve, por exemplo, é composto por uma tabela dimensão que pode ter relacionamentos com outras tabelas dimensão, além da tabela fato. Essa estrutura normalizada em níveis hierárquicos das tabelas pode aumentar a eficiência no armazenamento de dados, mas pode também tornar as consultas mais lentas e complexas, o que pode ser um fator a ser considerado na escolha entre os modelos, conforme destacado por Fana *et al.* (2021). Uma breve comparação entre os esquemas demonstra suas principais diferenças e permite uma escolha mais adequada para cada situação (Apêndice B, Quadro B. 6).

Considerando as características, integridade e relacionamento dos dados disponíveis, os requisitos demandados, a tecnologia utilizada para o desenvolvimento da solução que foi o *Power BI* e o objetivo deste trabalho, optou-se pela adoção do esquema misto utilizando ambos, tanto modelo *floco de neve* como o modelo estrela.

De acordo com Hassan *et al.* (2021), as tabelas dimensão definem o nível de detalhe que se deseja dos dados (filtros) e as tabelas fato definem os atributos de medida (métricas e ou indicadores-chave de desempenho: em inglês, KPIs) a partir dos filtros definidos pelo usuário. A utilização do esquema de modelagem dimensional *floco de neve* em um *Data Warehouse* possibilitou a organização eficiente das tabelas dimensão e fato dos três módulos que foram desenvolvidos para o MVP proposto.

Segundo Kimball *et al.* (2008), os atributos podem ser classificados em quatro categorias distintas. Primeiro, temos os atributos hierárquicos que permitem a organização dos dados em uma estrutura hierárquica, exemplificada por níveis como região, estado e cidade. Em segundo lugar, encontram-se os atributos descritivos, que fornecem um detalhamento sobre as características de um objeto específico, incluindo nome, descrição, código e categoria. A terceira categoria é constituída pelos atributos temporais, que estão relacionados à demarcação temporal e englobam data, hora, semana, mês, trimestre e ano. Finalmente, os atributos de medida são os que representam as métricas ou indicadores associados a um objeto, tais como quantidade, peso, preço e

medida. Cada uma dessas categorias tem um papel crucial na estruturação e análise de dados, contribuindo para a compreensão integral do objeto em estudo.

Os indicadores-chave de desempenho (KPIs) são medidas quantitativas que indicam a performance de uma instituição em relação aos seus objetivos e metas estabelecidos. Segundo alguns autores, KPIs são importantes para avaliar o desempenho de uma organização em relação a sua missão, visão e objetivos estratégicos (Albats *et al.*, 2018; Singh *et al.*, 2021; Skyrius, 2021).

A utilização de KPIs tem sido amplamente difundida em diferentes áreas, incluindo a gestão da educação e instituições de ensino superior (Albats *et al.*, 2018; Denwattana *et al.*, 2016; Hoffmann, 2021). No contexto da gestão educacional, os KPIs podem ser usados para avaliar a eficácia das políticas e estratégias adotadas e monitorar a qualidade do ensino e dos serviços prestados pela instituição (Guitart *et al.*, 2015).

Além disso, os KPIs podem ser customizados de acordo com as especificidades de cada ICTs, o que possibilita uma análise mais precisa e personalizada dos resultados (Albats *et al.*, 2018; Rabova *et al.*, 2014), a partir dos filtros selecionados dinamicamente pelo usuário final. Dessa forma, a utilização de KPIs pode contribuir significativamente para a tomada de decisões embasadas em dados concretos. Definida a modelagem e como os KPIs serão elaborados (Subseção 3.5.7), a seguir, é apresentada a concepção e a estrutura básica para a criação dos painéis visuais (*dashboards*)<sup>6</sup>.

De acordo com Skyrius (2021), os *dashboards* devem conter métricas e KPIs (Subseção 3.5.8) que representem aspectos relevantes sobre a atividade analisada, além de aspectos complementares que reflitam o contexto ou possíveis tendências. Para a construção do *dashboard* optou-se por buscar uma solução pronta no mercado que pudesse ser adaptada às necessidades do projeto, uma vez que a criação de um sistema próprio demandaria um investimento financeiro e de tempo consideráveis. Dentre as soluções disponíveis, a plataforma *Power BI* foi a escolhida, já que atende às demandas deste do projeto, é gratuita, oferece um recurso de perguntas e respostas e é uma ferramenta comumente utilizada em estudos acadêmicos ou já implementadas pelo governo brasileiro (Brasil, 2020; 2021a; 2022b), o que facilita futuras integrações desta solução proposta com as já existentes.

---

<sup>6</sup> Disponível em <https://prequest.websiteseuro.com/bi>



#### 2.4.3.4 *Processamento Analítico Online*

O Processamento Analítico Online (em inglês, *Online Analytical Processing* – OLAP), originário das consultas em bancos de dados, adaptou-se para abordar dimensões específicas de negócios, tais como tempo, local e produto. Ele é visto por especialistas tais quais Chaudhuri *et al.* (1997) como essencial na análise multidimensional de dados. Distintamente da *data mining*, que busca identificar padrões em conjuntos de dados, o OLAP oferece a capacidade de definir dimensões de negócios com clareza e precisão. Essa definição clara permite projetar questões comerciais sobre os dados e, assim, fornecer *insights* mais compreensíveis para os usuários (Skyrius, 2021).

Enquanto a *data mining* se concentra na descoberta de novos padrões, as consultas OLAP utilizam padrões já estabelecidos (Sharon *et al.*, 2022). O OLAP não apenas facilita o acesso e a visualização de dados com base em dimensões claramente definidas, mas também permite aos usuários adaptar essa visualização, utilizando estruturas como o Modelo Floco de Neve, tornando o processo de análise altamente flexível (Kongthanasuwan *et al.*, 2023). Essa flexibilidade e precisão na análise são cruciais, sobretudo, quando se compara a adoção do OLAP em relação a outras abordagens, como a *data mining*, que, por vezes, podem ser ambíguas para os usuários finais.

#### 2.4.3.5 *Data Mining*

A Mineração de Dados (em inglês, *Data Mining* - DM), objetiva identificar padrões em vastos conjuntos de dados, convertendo-os em conhecimento (Apêndice C). Como complemento, o Apêndice C apresenta uma análise multivariada de agrupamento, além do desenvolvimento de um índice, ambos elaborados utilizando o pacote R. Essas análises foram necessárias porque a ferramenta adotada, o Power BI, não oferecia suporte para tais procedimentos.

Como segmento avançado do BI, a DM busca informações, relações e padrões ocultos através de múltiplos processos, incluindo aprendizado de máquina e estatística. Esses possibilitam discernir *insights*, tais como segmentação de clientes, recomendações customizadas e previsões de comportamentos futuros (Nalawade *et al.*, 2021; Ozkaya *et al.*, 2021).

O propósito central de DM consiste em aplicar algoritmos para descobrir padrões úteis, utilizando ferramentas estatísticas (Schindler, 2016; Steiner *et al.*, 2006; Zhang *et*

*al.*, 2018). No contexto da estatística e Análise Exploratória de Dados (em inglês, EDA), a EDA é vital para o reconhecimento desses padrões e, frequentemente, se associa aos métodos de interdependência multivariada (Fávero *et al.*, 2020; Hair *et al.*, 2019).

Diversos processos de DM podem ser empregados para a identificação de padrões, entre elas: associação, classificação, agrupamento, previsão, padrões sequenciais e discriminação (Han *et al.*, 2012; Nalawade *et al.*, 2021).

Neste estudo, selecionou-se a Análise de Agrupamentos (em inglês, *Cluster Analysis* – CA) (Apêndice C.1) e a Análise Fatorial Exploratória ou *Exploratory Factor Analysis* (EFA) por Componentes Principais (em inglês, *Principal Components* - PCA) (Apêndice C.2), cujos resultados podem ser expressos em relatórios estatísticos ou painéis visuais ou *dashboards*.

Os métodos de interdependência multivariada, aplicados em múltiplos campos do saber, estendem suas funcionalidades a áreas diversas, como a classificação de aeronaves em aviação (Santos, E. *et al.*, 2019), detecção de agrupamentos de tumores em ressonâncias (Kumar *et al.*, 2022), seleção de características do gás natural urbano (Dashdondov *et al.*, 2022) e o uso do agrupamento hierárquico e análise fatorial exploratória por componentes principais para determinar as principais ICTs e variáveis, respectivamente às análises realizadas, considerando um conjunto de dados multidimensionais, para visualizar as relações de TH entre ICTs-GOV-SPP (Andrade *et al.*, 2022). Esses métodos também são pertinentes para entender a mensuração e análise de pesquisa, desenvolvimento tecnológico e inovação, enquadrando-se tanto na tese do modelo TH de inovação (Etzkowitz *et al.*, 1995; 2000) quanto na pesquisa do SNI (Dore, 1988; Metcalfe, 1995; Nelson, 1993).

#### 2.4.3.6 Consultas e Relatórios

Os sistemas de BI permitem que os usuários criem consultas e/ou relatórios personalizados com base em suas necessidades específicas. Essas consultas, como apontado por Srivastava *et al.* (2022), permitem que o usuário gere e entregue dados relevantes do servidor BD da instituição e tendem a ter um propósito de uso único, sem a capacidade de serem incorporadas em processos de negócios coletados no futuro. Além disso, podem ser executadas nos BD ou DW.

Há também os relatórios, incluindo os *ad hoc* mencionados por Srivastava *et al.* (2022) que podem reunir todos os relatórios da organização independentemente da solução da TI proposta.

As consultas e relatórios podem ser visualizados em diversos formatos, como tabelas, gráficos, *dashboards*, planilhas e documentos texto. Essas informações são cruciais para acompanhar ou monitorar o desempenho dos “negócios”, avaliar resultados e identificar áreas de melhoria, dentre outras aplicações.

#### 2.4.3.7 *Visualização de Dados*

Os painéis ou *dashboards* e *scorecards*, enquanto interfaces visuais, permitem visualizar o desempenho do negócio em tempo real e de maneira estratégica. Enquanto o primeiro costuma oferecer atualizações em tempo real, o segundo proporciona uma visão mais condensada e direcionada. Essas ferramentas emergiram nos anos 70 como suporte aos sistemas de apoio à decisão (Mould *et al.*, 2016), objetivam facilitar a compreensão dos extensos volumes de dados corporativos, otimizando o tempo de acesso à informação (Stair *et al.*, 2010).

A crescente aplicação do *dashboard* no ecossistema das ICTs evidencia sua importância, conforme corroborado por diversos estudos (Mould *et al.*, 2016; Stewart *et al.*, 2022; Wang *et al.*, 2015). Este interesse deriva da capacidade dessa ferramenta em auxiliar gestores a interpretar grandes volumes de dados, monitorar indicadores-chave de desempenho (KPIs) e identificar tendências (Rabova *et al.*, 2014). Acrescente-se a essa relação o atendimento às demandas dos órgãos públicos de controle em tempo real. Diante disso, evidencia-se um desafio: a necessidade de soluções tecnológicas que suportem decisões no planejamento, formulação implementação e monitoramento de políticas voltadas aos canais de PD&S.

#### 2.4.3.8 *Integração do UX Design em Dashboards*

A eficácia dos *dashboards* reside em sua apresentação clara e organizada, e a necessidade de *UX design* torna-se crucial neste e demais contextos (Erds *et al.*, 2019). Tal *design* considera elementos como cor, contraste e tipografia, visando criar interações visuais intuitivas com os usuários (El-Gammal *et al.*, 2022). Martins *et al.* (2022) exemplificam a integração bem-sucedida do *UX design* em *dashboards*, estabelecendo

um equilíbrio entre componentes gráficos e funcionais, útil para atender demandas específicas, como as das ICTs.

#### 2.4.3.9 Importância dos Indicadores-Chave de Desempenho

Os Indicadores-Chave de Desempenho (KPIs), enquanto métricas quantificáveis, são empregados para medir e monitorar o desempenho institucional (Albats *et al.*, 2018; Rabova *et al.*, 2014). Existem várias categorias de KPIs, sendo as principais baseadas em tempo, custo e qualidade (Bertagnolli, 2022). Esses indicadores, de acordo com a ABNT (2000), facilitam o monitoramento contínuo das metas e estratégias, evidenciando o progresso e os resultados, baseado em evidências pelos gestores (Skyrius, 2021).

#### 2.4.3.10 Balanced Scorecard e Seu Impacto na Estratégia

O *Balanced Scorecard* (BSC), conforme delineado por Kaplan *et al.* (1997, p. 2), traduz missões e estratégias organizacionais em medidas abrangentes de desempenho, segmentadas em quatro perspectivas principais. Essas incluem: perspectiva financeira, perspectiva do cliente, perspectiva dos processos internos e perspectiva da aprendizagem e crescimento (Kaplan, 1992; Kaplan *et al.*, 1997; Niven, 2007). A capacidade de alinhar essas perspectivas aos interesses dos tomadores de decisão, gestores ou *stakeholders* torna o BSC uma ferramenta estratégica adaptável, como exemplificado por modelos derivados, como o *Public Value Scorecard* (PVS) (Azevedo, 2012).

### 2.4.4 Interações entre BI e a Revisão da Literatura

A técnica de BI tem sido uma ferramenta valiosa para visualizar e entender os diversos canais dos ecossistemas das ICTs ao longo dos anos. Como destacado por Santos, T. R. *et al.* (2021) e Santi *et al.* (2018), a implementação e o uso de BI não se restringem apenas à visualização, mas também para suportar decisões administrativas e acadêmicas e para melhorar processos em diversos canais desses ecossistemas.

Além disso, ao combinar a técnica de BI com a tecnologia de DW, há um aumento significativo em sua relevância em estudos tanto no Brasil quanto globalmente. Essa combinação possibilita um tratamento de dados mais profundo. Esse tratamento

possibilita identificar tendências e padrões essenciais para a tomada de decisão, conforme evidenciado por Santos, T. R. *et al.* (2021).

No Brasil, em órgãos governamentais como MCTI e MEC, e em empresas como a *StelaTek*, observa-se a crescente adoção de soluções tecnológicas que empregam BI e *dashboards* em ecossistemas de ICTs. Exemplos incluem a plataforma Painel Universidade 360, o Observatório Nacional de Ciência, Tecnologia e Inovação, o Portal Integra e a solução privada Stela Experta®, desenvolvida pelo SPP. Enquanto as plataformas governamentais focam na análise e monitoramento de dados relacionados ao ensino, administração e finanças, a solução privada, a StelaExperta©, como aborda primordialmente os canais de PD&I e CT&I.

Este trabalho de pesquisa propõe o desenvolvimento de uma tecnologia que, além dos já mencionados, incorpora canais de Serviços. Isso potencialmente amplia as capacidades de análise e monitoramento das plataformas existentes, principalmente, as públicas. Entretanto, como ressaltado por Musa *et al.* (2019), a eficácia na implementação dessas soluções depende de uma gestão tecnológica adequada e da inclusão de métricas relevantes. Corroborando com a perspectiva acima, Santos (2010) salienta “ a necessidade da RFEPCCT em investir no desenvolvimento ou aquisição de plataformas especializadas que facilitassem o acesso rápido e aberto do conhecimento produzido”.

Um resumo das plataformas, tanto públicas quanto privadas, que aplicam técnicas de BI em conjunto com *dashboards* no ecossistema de ICTs, atualmente focam em monitorar e analisar dados que auxiliam na tomada de decisões. Estas decisões estão relacionadas aos canais de ensino, administração, finanças, indicadores acadêmicos, serviços de inteligência, infraestrutura e desenvolvimento tecnológico (Apêndice B, Quadro B. 1).

Guitart *et al.* (2015) e Wang *et al.* (2013) complementam a visão oferecida por Wang (2015), sublinhando a eficácia das soluções baseadas em BI no aprimoramento da tomada de decisão e na otimização da gestão em ICTs. Eles destacam a necessidade de uma análise mais profunda para identificar lacunas e direções futuras no campo da BI, especialmente em ambientes de ICTs.

Medina *et al.* (2018) e Damyanov *et al.* (2019) destacam que as soluções tecnológicas baseadas na técnica de BI podem aprimorar a gestão de dados e informações em ICTs. Isso corrobora com a perspectiva de Musa *et al.* (2019) sobre o impacto da gestão adequada na implementação bem-sucedida de sistemas baseados em BI.

A revisão da literatura também permitiu identificar estudos que aplicaram a técnica de BI associada à tecnologia do *dashboards* e DW no contexto das ICTs, no Brasil (Azevedo *et al.*, 2021; Hoffmann, 2021; Santos, Thiago Rizzi *et al.*, 2021; Zhai *et al.*, 2010) e no mundo (Bianchi *et al.*, 2022) por exemplo, identificou apenas três estudos (Cahyadi *et al.*, 2015; Castillo *et al.*, 2018; Duro Novoa *et al.*, 2016) que adotaram o conjunto de soluções baseada em BI, também adotado nesta tese, para investigar os canais de P&D. Ao revisar o trabalho de Bianchi *et al.* (2022), referente aos 89 trabalhos relacionados a aplicação da técnica de BI associadas as tecnologias de *dashboards* e DW no contexto das ICTs, apenas três foram identificados abordando esse contexto.

Adicionalmente, durante a revisão, foram identificados mais dois estudos, sendo um deles desenvolvido por Denwattana *et al.* (2016), e outro por Azevedo *et al.* (2021), sendo esse último envolvendo uma ICTs pública brasileira. Destaca-se que nenhum desses estudos abordou os canais investigados no contexto das interações TH, apenas consideram o ambiente interno das ICTs, restringindo a visualização e organização dos dados para tomadas de decisões endógenas.

De modo geral, os estudos brasileiros ou estrangeiros propuseram arquiteturas, protótipos de *dashboards* e modelos de projeto para a construção de DW (Bassil, 2012; Castillo *et al.*, 2018) como suporte a BI e ao *dashboard*. Esses estudos limitam-se, em geral, a considerar métricas como, por exemplo, patentes, projetos de pesquisa e publicações de maneira isolada ou associada. Contudo, métricas em outras áreas acadêmicas, como gestão, finanças e ensino, são mais frequentes.

Em relação aos SPP, foi proposta uma estrutura de integração e análise de dados de patentes com base em BI (Zhai *et al.*, 2010). Também foi sugerida uma arquitetura de BI para monitorar e analisar indicadores nacionais de P&D (Khatibi *et al.*, 2017).

Wang *et al.* (2013) propuseram uma plataforma de serviços de informação para ICTs, utilizando tecnologia DW, a fim de organizar e atender às demandas das interações ICTs, SPP e GOV relacionada à prestação de serviços acadêmicos, científicos e/ou tecnológicos. Di Tria *et al.* (2015) apresentaram uma arquitetura de BI e DW para ICTs com foco na avaliação da pesquisa e didática. Fernandez *et al.* (2017) desenvolveram um modelo de qualidade aplicável a uma solução de BI para ICTs baseado na série de normas ISO/IEC 25000 (Standardization *et al.*, 2011).

Outro estudo foi o de Hoffmann (2021), que apresentou um caso de aplicação de BI e *dashboard* e DW em uma ICTs federal brasileira. Eles mostraram como o uso dessas ferramentas permitiu demonstrarem as diferenças relevantes entre os *campi* da ICTs e

entre o perfil dos servidores em diversos aspectos e cenários. Eles destacam que, embora a pesquisa apresentada tenha contribuições relevantes para a gestão, faz-se necessário considerar as limitações e possibilidades que o uso de BI e *dashboards* e DW oferecem dentro do contexto mais amplo da TH, da TAN e da TVD.

A aplicação da técnica de BI, aliada às tecnologias de DW, *dashboards* e Teorias de Tecnologia ou de Gestão da Informação, pode mitigar as dificuldades e limitações mencionadas nos estudos citados anteriormente. Assim, proporciona uma base metodológica sólida para tomada de decisões e monitoramento do desempenho das ICTs. Essas tecnologias podem auxiliar usuários, em especial, os gestores, na identificação de oportunidades e desafios, na alocação eficiente de recursos e na formulação de políticas mais assertivas (Rabova *et al.*, 2014), possibilitando um ambiente propício ao desenvolvimento inovador e ao crescimento socioeconômico.

Além disso, a técnica de BI e as tecnologias a ela associadas podem melhorar a comunicação e colaboração entre os atores da TH, facilitando a disseminação de informações e conhecimentos, promovendo maior cooperação e alinhamento de interesses (Mould *et al.*, 2016; Stewart *et al.*, 2022). Isso pode levar a um ambiente mais propício à inovação e ao desenvolvimento de soluções tecnológicas e científicas (Albats *et al.*, 2018; Lino, 2013).

Portanto, a utilização de BI nas ICTs, corrobora para integrar e otimizar os dados relacionados aos diversos canais relacionados as interações teoria da Tripla Hélice (TH), especificamente nesse estudo considerou-se os dados relacionados aos canais de PD&S, reforçando a colaboração entre seus atores, incentivando a inovação e o avanço socioeconômico. Essa conexão é particularmente evidente quando se consideram as variáveis pesquisa, publicações científicas, investimentos em P&D, pesquisadores, patentes, consultorias, dentre outras. Essas variáveis, frequentemente investigadas à luz da teoria da TH, representam aspectos fundamentais dos ecossistemas de inovação.

Conforme constatado durante a revisão da literatura, vários são os métodos para analisar as interações a partir da teoria TH, porém estudos incorporando soluções de BI anda são escassos. Foi identificado que a maioria deles, apesar de analisarem elementos específicos, carecem de uma padronização mínima que permita sua utilização como elementos comparativos em estudos *cross-country* ou entre ICTs. Outro ponto a destacar é que, embora muitos trabalhos adotem a análise quantitativa, tais trabalhos utilizaram pequenas amostras de dados, o que dificulta a compreensão macro dos resultados, sendo estes aplicáveis apenas aos seus respectivos estudos de caso.

Em conclusão, assim como a revisão de Santos, Thiago Rizzi *et al.* (2021) aponta, o processo de descoberta de conhecimento usando dados é essencial. A integração, filtragem, gestão e monitoramento desses dados, especialmente em relação aos canais de PD&S em ICTs, não são apenas necessários, mas também representam um campo em rápido crescimento que requer pesquisa integrada e colaborativa.



### 3 Metodologia

Antes de apresentar este capítulo, é importante ressaltar que este trabalho de pesquisa, aqui proposto, tem como objetivo apresentar uma solução tecnológica, cujo registro de *software* nº BR5120230010197, foi concedido sob a forma MVP<sup>7</sup> de uma plataforma de serviço *online*, baseada em BI<sup>8</sup>, para gestão dos canais de PD&S<sup>9</sup>, em ICTs<sup>10</sup>, fundamentada nas teorias da Análise de Negócios, de Visualização de Dados e da Tripla Hélice da Inovação.

A metodologia deste trabalho consistiu em uma abordagem sistemática e detalhada do trabalho proposto. Ela focou na fundamentação teórica, abordagem de pesquisa, levantamento de requisitos e desenvolvimento inicial, incluindo a concepção do modelo de BI, com exceção de alguns detalhes técnicos dispostos nos apêndices. Com isso foi possível assegurar uma base sólida necessária para a construção e validação da solução tecnológica proposta.

O presente trabalho é de natureza aplicada e utiliza o método empírico-dedutivo, com desenho explicativo, abordagem quantitativa e longitudinal, pois coleta dados em dois ou mais momentos do tempo (Gray, 2016) e o estudo de caso (Yin, 2015) adotado refere-se à RFEPCT<sup>11</sup>.

Destaca-se que a coleta de dados foi conduzida no período de 1º de janeiro de 2008 a 31 de dezembro de 2018, abrangendo um total de 28 variáveis analisadas neste estudo, das quais 23 são discretas, 2 contínuas e 3 nominais.

A metodologia proposta para o desenvolvimento da solução inclui o contexto do estudo, a revisão bibliográfica, a identificação de fontes de dados, o desenvolvimento do modelo de dados, o desenvolvimento da solução, a validação da solução e a proposição do MVP<sup>12</sup>. Na próxima página um diagrama do esquema metodológico encontra-se apresentado no Diagrama 1, que está de acordo com alguns dos subitens do presente capítulo, bem como com a estrutura da tese.

---

<sup>7</sup> Mínimo produto viável.

<sup>8</sup> *Business Intelligence*.

<sup>9</sup> Pesquisa, Desenvolvimento e Prestação de Serviços.

<sup>10</sup> Relacionadas às Instituições de Ensino, Pesquisa, Ciência e Tecnologia.

<sup>11</sup> Rede Federal de Educação Profissional, Científica e Tecnológica.

<sup>12</sup> Disponível em: <https://prequest.websiteseuro.com/bi>

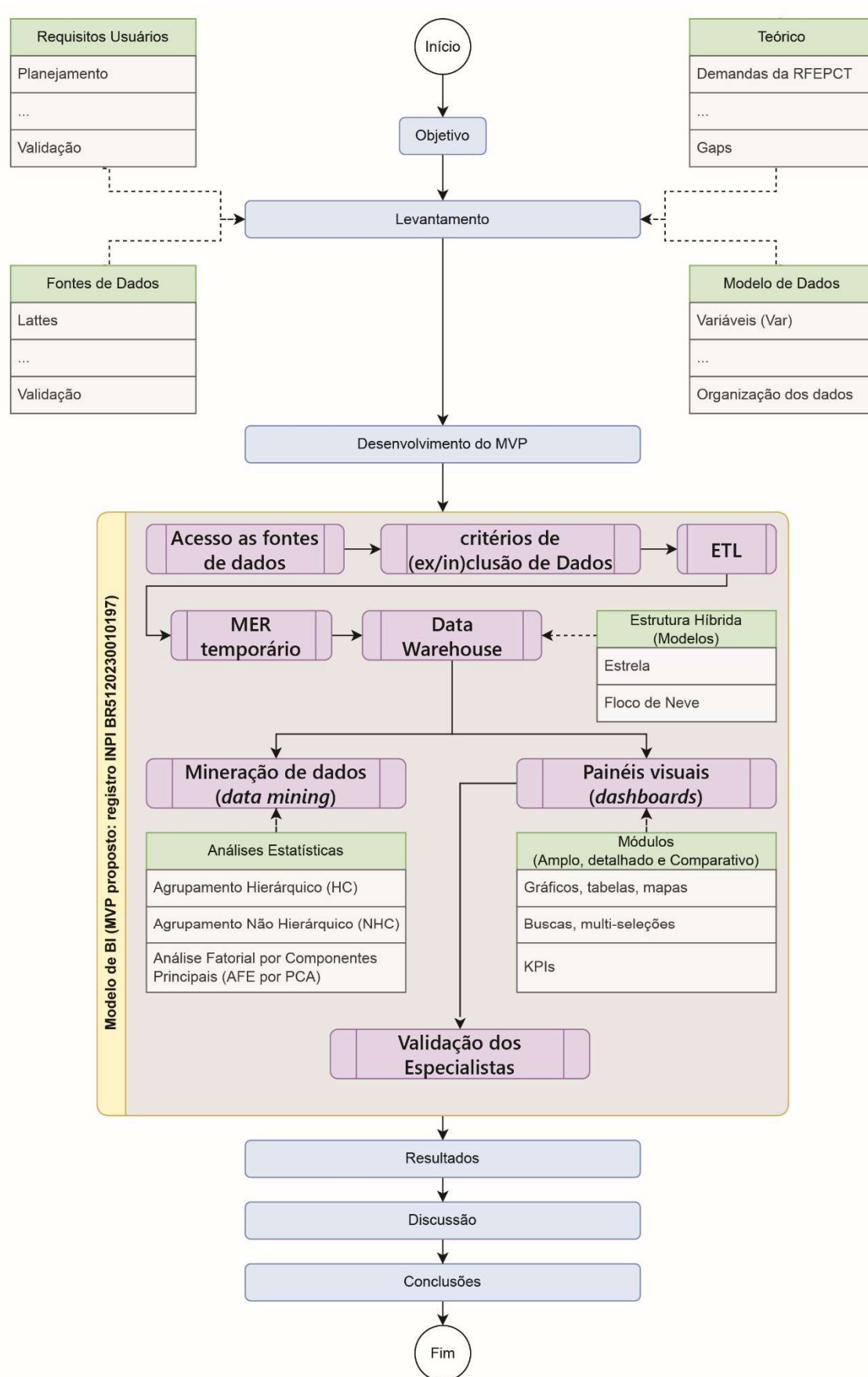


Diagrama 1: Esquema metodológico aplicado nesta tese.

A adoção dessa solução pode contribuir significativamente para o avanço do conhecimento dos canais de PD&S num ecossistema das ICTs, fornecendo subsídios práticos à tomada de decisão, investigação, gestão e monitoramento embasados em evidências. A seguir, são apresentados mais detalhes sobre cada uma dessas etapas.

Tornou-se essencial preservar e ocultar detalhes específicos, como partes do processo de ETL e da programação em linguagem Dax no PowerBI, que permitem a seleção multivariada dos dados, para proteger o segredo industrial e o sigilo comercial durante negociações com possíveis investidores em segmentos metodológicos deste estudo. O potencial comercial da solução tecnológica proposta foi identificado pelo grupo de pesquisa denominado de Núcleo Multidisciplinar de Desenvolvimento e Inovação Tecnológica (NUMDI). Tal motivação levou o proponente deste trabalho, em conjunto com o NUMDI, o qual, seus integrantes tiveram como referência a experiência prática na colaboração e gestão de uma Pró Reitoria de Pesquisa e Inovação da RFEPCCT e participação nos Fóruns dos Pró Reitores de Pesquisa e Inovação (FORPOG), entre 2010 a 2017, inserido em um ambiente de trabalho que suscitava a necessidade de implementação por soluções como a proposta neste trabalhos, implementação esta corroborada pelos especialistas avaliadores do MVP aqui proposto.

### **3.1 Seleção e Análise dos Estudos Utilizados neste Trabalho**

Neste trabalho, buscou-se coletar informações relevantes por meio da seleção e análise de estudos. As características do ecossistema da RFEPCCT, a abordagem dos seus canais de PD&S, a identificação de fatores de CT&I e PD&I influentes, a teoria adequada para o entendimento do fenômeno e as técnicas e tecnologias usadas para a resolução de problemas similares foram os tópicos considerados. Para tal, recorreu-se a bases de dados acadêmicas e científicas, artigos, livros e outras publicações especializadas. As bases científicas consultadas foram o Portal Periódicos Capes, a *Web of Science* e a *Scopus*.

Nos estudos publicados, a inclusão foi considerada se tivessem passado por revisão por pares, disponibilidade nas bases científicas consultadas, escritos em português, inglês ou espanhol, e publicados preferencialmente entre 2008 e janeiro de 2023. Contudo, quando necessário, artigos seminais, históricos ou outras fontes de relevância foram incluídos. Para a exclusão, considerou-se: título nulo, resumo nulo, autores nulos e a impossibilidade de acesso *online* (acesso pago ou disponível apenas em bibliotecas presenciais) ou ao arquivo em formato de pdf ou físico.

As buscas (para detalhes sobre o processo, critérios e *springs* de busca consultar Apêndice B, Quadro B. 7 a B. 11) focaram preferencialmente nos *abstracts*, títulos e palavras-chave, nesta ordem de relevância. Optou-se pela utilização do gerenciador bibliométrico *Endnote V20*<sup>®</sup> e a biblioteca *Bibliometrix* para a análise bibliométrica dos artigos, desenvolvida em linguagem *R*.

A seleção e a análise dos estudos utilizados neste trabalho foram conduzidas através de um processo rigoroso e sistemático, que compreendeu as seguintes etapas:

- i. **Definição de critérios de inclusão e exclusão:** foram estabelecidos os critérios claros e objetivos para determinar quais artigos seriam relevantes e adequados para a nossa pesquisa.
- ii. **Busca nas bases de dados:** foram realizadas buscas abrangentes em bases de dados Portal Periódicos Capes, a *Web of Science* e a *Scopus*, utilizando palavras-chave e termos relacionados ao tema de pesquisa.
- iii. **Triagem inicial dos artigos:** com base nos critérios de inclusão e exclusão, foram filtrados os artigos identificados na busca, excluindo aqueles que não atendiam aos critérios estabelecidos.
- iv. **Análise de qualidade e relevância:** Foram avaliados a qualidade metodológica — se, por exemplo, apresentavam as informações como abrangência do estudo, variáveis utilizadas, etc. — e a relevância dos artigos pré-selecionados. Para isso, foram utilizadas ferramentas como EndNote e NVivo, e critérios específicos, como por exemplo estarem classificados entre os quartis 1 e 2, de acordo com o JCR, ou Qualis Capes A1 até B2, para determinar quais artigos seriam incluídos na análise final.
- v. **Extração e síntese dos dados:** foram identificados os dados e as variáveis referentes aos canais de PD&S utilizadas nos artigos selecionados. Esses foram sintetizados no Quadro B. 5 (Apêndice B), que apresenta os principais resultados de forma organizada e compreensível.

A seção 2 buscou organizar as principais teorias, métodos, técnicas e tecnologias frequentemente utilizadas para explorar, analisar e apresentar os dados relativos aos canais de PD&S. As etapas de 1 a 5 foram executadas separadamente e descrevem os critérios adotados para a seleção dos artigos analisados e uma síntese dos resultados obtidos para cada tópico abordado no referencial teórico.

A etapa 1 (Apêndice B, Quadro B. 8) sintetiza estudos empregados na investigação da evolução temporal dos canais de PD&S do ecossistema da RFEPCT. A subseção 2.1 focou em revisar os estudos já realizados sobre a RFEPCT buscando identificar os elementos frequentemente relatados, variáveis consideradas, soluções desenvolvidas, resultados identificados e potenciais lacunas que, ao serem preenchidas, possam aprimorar ou assegurar a interação TH com as ICTs.

A etapa 2 (Apêndice B, Quadro B. 9) apresenta a síntese dos trabalhos utilizados para contextualizar alguns fatores nacionais de CT&I e PD&S. A Subseção 2.2 revisou alguns dos fatores nacionais de CT&I e PD&S em que o ecossistema das ICTs se desenvolveu nas duas últimas décadas, em especial de 2008 a 2018. Para construí-lo utilizou-se como principal base de dados o Portal Periódicos Capes, *Web of Science* e *Scopus* e o diretório de dados brutos da plataforma Sucupira e Tribunal de Contas da União.

A etapa 3 (Apêndice B, Quadro B. 10) sintetiza a adoção da teoria TH na investigação dos fenômenos relacionados aos canais de PD&S. A subseção 2.3 explora técnicas e variáveis frequentemente implementadas por essa teoria. Além disso, abordou-se a aplicação de técnicas de BI, aliadas a tecnologias de DW e *dashboard*, em ICTs, visando identificar canais, variáveis, o número de fontes de dados consideradas e as técnicas e tecnologias utilizadas.

A etapa 4 (Apêndice B, Quadro B. 11) consolida pesquisas que associam a técnica de BI com as tecnologias DW e *dashboard*, no contexto das ICTs, sob o arcabouço da teoria da TH, TAN e TVD. A Subseção 2.4 oferece uma revisão da implementação dessas técnicas e tecnologias, visando identificar os canais de PD&S, previamente explorados, seja de maneira isolada ou combinada, bem como as ferramentas utilizadas para tal discernimento ao longo do tempo.

A etapa 5 sintetiza as fontes utilizadas para traçar a história da Educação Profissional, Científica e Tecnológica (EPCT) brasileira (Apêndice B, Quadro B. 7). Esta etapa apresenta os critérios que levaram a selecionar os dados da RFEPCT como estudo de caso. A investigação aborda a evolução da EPCT, desde decretos e leis iniciais do século XX, que instituíram escolas de aprendizes artífices e um ensino profissional primário e gratuito, até as normativas contemporâneas que estabeleceram a RFEPCT. A expansão e transformação das intuições vinculadas a essa formação e sua interação com o modelo TH são também exploradas, provendo um panorama da dinâmica da inovação

e estatísticas correlatas à RFEPCT. Estes estudos forneceram uma base para a análise temporal da evolução da educação profissional brasileira.

### **3.2 Levantamento de Requisitos dos Usuários**

A construção do DW, respeitando os fundamentos da TAN (Subseção 2.4.2.3), visa o desenvolvimento de uma solução Web baseada em BI. Essa solução pretende permitir uma análise objetiva e confiável dos canais de PD&S para ecossistemas de ICTs. Para alcançar o objetivo proposto um levantamento das necessidades e requisitos dos usuários foi conduzido a partir de reuniões virtuais entre o pesquisador deste estudo com alguns gestores da RFEPCT e de uma Universidade Federal, os quais já assumiram o cargo de Pró-reitores de Pesquisa e Inovação em suas respectivas ICTs.

Os requisitos identificados inicialmente foram:

- **Integração de Dados:** Facilitar a consolidação de dados de diferentes fontes para proporcionar uma visão unificada.
- **Análise e Visualização:** Permitir análises complexas e visualização intuitiva dos dados para auxiliar na tomada de decisão.
- **Acesso e Transparência:** Garantir fácil acesso aos dados e promover a transparência nas atividades das ICTs.
- **Resposta às Demandas:** Capacidade de responder rapidamente às demandas dos setores produtivos, governo e sociedade.
- **Gestão do Conhecimento:** Fomentar a criação, compartilhamento e aplicação do conhecimento gerado.

As entidades identificadas inicialmente foram:

- **ICTs:** Instituições que realizam pesquisa, desenvolvimento e oferecem serviços tecnológicos.
  - **Atributos:** Nome, localização, tipo (pública), pertencente a RFEPCT.
- **Projetos de PD&S:** Projetos realizados nas áreas de pesquisa, desenvolvimento tecnológico e prestação de serviços.
  - **Atributos:** ID do projeto, título, área temática, financiamento, status, resultados.
- **Usuários:** Incluem gestores das ICTs, pesquisadores, governo, setores produtivos e sociedade.

- Atributos: interesse, demandas, indicadores.

Os Relacionamentos identificados inicialmente foram:

- ICTs e canais de PD&S: Uma ICT pode entregar diversos resultados relacionais aos canais de PD&S.
- Canais de PD&S e Usuários: os dados devem atender às demandas específicas dos usuários durante a análise de um determinado cenário selecionado por ele.
- ICTs e Usuários: ICTs devem interagir diretamente com os usuários, permitindo que os usuários identifiquem soluções para suas demandas e que as ICTs compartilhem seus resultados com os usuários de forma transparente.

Durante o projeto outros requisitos, entidades, atributos e relacionamentos foram identificados, tais como:

- Integração com sistemas existentes:** Os usuários requerem a integração do DW com sistemas preexistentes e novos, tais como a Plataforma Lattes, o DGP, o INPI, entre outros. Isso significa que o DW precisa ser capaz de se conectar com diversas bases de dados institucionais e sistemas externos, além de fontes de dados governamentais relevantes para a análise dos canais de PD&S.
- Segmentação, análise e relatórios de PD&S para Auditorias:** disponibilização de uma visualização interativa e consolidada dos indicadores relacionados aos canais de PD&S de suas ICTs, relativos a CT&I e PD&I, que possa ser consultada por gestores e atenda às solicitações de órgãos como Ministério Público Federal (MPF), TCU e CGU.
- Segmentação, análise e relatórios do perfil dos pesquisadores:** isso inclui informações sobre suas formações, produção científica, projetos de pesquisa, orientações, colaborações com setores produtivos públicos e privados, patentes registradas, dentre outros aspectos relevantes para avaliar a produtividade e contribuição.
- Segmentação, análise e relatórios dos indicadores de PD&S:** consultar e monitorar os indicadores de PD&S, incluindo informações sobre o

número de artigos completos publicados em periódicos, projetos de pesquisa em andamento, orientações de pesquisa concluídas, patentes concedidas e projetos de extensão desenvolvidos etc. Além disso, o *dashboard* deve possibilitar a segmentação dos dados por áreas de concentração dos grupos de pesquisa, tipos de financiamentos ou investimentos recebidos e localização administrativa das ICTs.

- v) **Segmentação, análise e relatórios de desempenho das PD&S em cooperação/interação com os SPP:** deve disponibilizar uma consulta que apresente a quantidade e as características de pesquisas, projetos de desenvolvimento tecnológicos, produto, e/ou serviços em cooperação com SPP. Esse requisito possibilitará aos gestores das ICTs terem uma visão clara e objetiva das colaborações entre sua IECT-SPP como de outras para fins comparativos.
- vi) **Segmentação, análise e relatórios dos investimentos em PD&S:** deve disponibilizar um painel de controle que apresente de forma clara e detalhada os investimentos públicos e privados em PD&S realizados pelas ICTs, ao longo dos anos. Esse painel deve permitir que os gestores possam visualizar e analisar de forma objetiva e confiável os valores destinados aos diferentes tipos de projetos de pesquisa, quantidade de pesquisadores ativos, número de patentes registradas, entre outros indicadores relevantes. Além disso, a solução deve possibilitar a segmentação dos investimentos por fontes de financiamento (públicas e privadas), contribuindo para uma melhor compreensão do cenário financeiro e a tomada de decisões estratégicas.
- vii) **Painéis de PD&S para consulta pública:** esses painéis devem permitir que os usuários (gestores, pesquisadores, SPP, GOV ou sociedade civil) realizem consultas personalizadas, filtrem os dados por diferentes variáveis, como anos, regiões, áreas de pesquisa, entre outros, e obtenham *insights* relevantes de forma rápida e intuitiva.

### 3.3 Desenvolvimento do Modelo de Dados

A ausência de uma solução pública e confiável para a obtenção de dados precisos e objetivos sobre os canais de PD&S tem impactado a capacidade dos tomadores de



decisão, pesquisadores e sociedade em conhecer, filtrar, gerenciar e monitorar as informações geradas pelos dados. Esse contexto vem dificultado a análise e monitoramento a longo prazo dos canais de PD&S, relacionados à complexa interação entre as ICTs, SPP e GOV.

A modelagem dos dados (Figura 2, Figura B. 1, Figura B. 2 e Apêndice D), o processo de ETL (Apêndice E, Apêndice B, Figura B. 3), a modelagem do MVP (Apêndice F e G, Figura B. 4 - Figura B. 12) e apresentação dos dados (Apêndice F, Figura B. 13 - Figura B. 17), que foram planejadas, respeitando os fundamentos da TVD (Subseção 2.4.2.5). Esta escolha metodológica viabilizou decisões robustas para desenvolver a solução tecnológica que atenda o contexto da inteligência de negócios contemporâneo, cuja precisão e objetividade da interface são essenciais para uma tomada de decisão, análise e gestão baseadas em evidências.

No contexto atual, modelos de dados confiáveis e robustos são cruciais. O trabalho propõe um modelo para analisar canais de PD&S em ICTs, integrando diversas fontes e utilizando o esquema misto, combinando floco de neve e estrela, para assegurar a integridade e relação dos dados.

O modelo de dados utilizado para a obtenção de informações precisas e objetivas considerou as principais variáveis envolvidas nos canais de PD&S, bem como as relações entre elas. Essa escolha foi baseada nas informações coletadas na revisão bibliográfica (Subseção 2), na identificação das 7 fontes de dados consideradas (Apêndice B, Quadro B. 12 e Quadro B. 13) pois são fontes públicas que armazenam dados das ICTs referente a diversos canais como por exemplo os de PD&S, que foram adicionadas para garantir a integridade e o relacionamento dos dados carregados das 28 variáveis selecionadas (Apêndice B, Quadro B. 14) após a revisão da literatura e do levantamento dos requisitos e tratadas por meio do processo de ETL (Subseção 3.5.3 e Apêndice B, Figura B. 3).

O modelo de dados proposto é apresentado em um Modelo de Entidade de Relacionamento (A Apêndices F e B; Subseção 3.5.4 e Figura 2, Figura B. 1, Figura B. 2). Foram utilizadas 14 tabelas principais, cada uma com sua própria chave primária e colunas específicas, conectadas por meio de chaves estrangeiras para evitar a redundância de informações e garantir a consistência dos dados analisados dos canais de PD&S em ICTs (Apêndice B, Figura B. 1).

A adoção do esquema misto (estrela e floco de neve) foi importante para a eficiência do armazenamento de informações no DW (Subseção 3.5.5) para atender aos pressupostos das TAN e TVD consideradas (Subseção 2.4.2.3 e 2.4.2.5), além de garantir

a precisão e a confiabilidade dos dados utilizados na análise. A ferramenta Power BI (Subseção 2.4.3.3) foi utilizada para o desenvolvimento da solução *online* proposta sob o formato de um MVP empregando a tecnologia de *dashboard* (Subseção 2.4.3.7).

### **3.4 Identificação de Fontes de Dados e Variáveis Identificados no Referencial Teórico Analisado**

Esta subseção concentra as informações resultantes das fontes de dados que contém as variáveis selecionadas (Apêndice B, Quadro B. 14), para o desenvolvimento desse projeto. Elas foram identificadas durante a revisão teórica detalhada (Seção 2) e analisados durante o procedimento ETL (Subseção 3.5.3 e Apêndice E).

Os quadros de dados no Apêndice B (Quadro B.13, Quadro B.14 e Quadro B.15) apresentam as fontes de dados oficiais brasileiras consideradas, Características das 28 variáveis/métricas consideradas para a coleta dos dados referentes aos canais de PD&S e a categorização dessas variáveis em relação canais de PD&S que cada uma está relacionada. Eles oferecem uma visão geral dos dados coletados servindo como referencial para futuras comparações. A tese também descreve as quatro etapas executadas no processo.

A etapa 1 consolida as fontes de dados públicos e os referenciais teóricos frequentemente aplicados em estudos de canais de PD&S das ICTs no âmbito da Teoria da TH (Seção 2 e Apêndice B, Quadro B. 12 e Quadro B. 13). No Brasil, destacam-se 4 fontes principais: a Plataforma Lattes, o DGP, o RGIE e o INPI. Adicionalmente, outros métodos de coleta, como questionários, grupos focais e análise documental, são referenciados, demonstrando a gama de estratégias empregadas.

A etapa 2 sintetiza os canais e variáveis comumente considerados na literatura científica, relacionados ao marco teórico da TH (Seção 2 e Apêndice B, Quadro B. 5, Quadro B. 14 e Quadro B. 15), abordando os canais de PD&S. Foram localizadas 43 variáveis distribuídas em 6 dimensões: Geográfica, institucionais, de pesquisa, Desenvolvimento Tecnológico, de P&D e de serviços. Destaca-se que nem todas as variáveis identificadas durante o referencial teórico (Apêndice B, Quadro B. 5) foram selecionadas para serem incluídas neste estudo. Para sua inclusão o critério era que representasse o máximo de variáveis possíveis por canal e também contribuísse para atender aos requisitos identificados os Quadro B. 14 e Quadro B. 15 (Apêndice B) apresentam as variáveis selecionadas e a quais canais ela está vinculada.

A etapa 3 apresenta um resumo das 7 fontes de dados utilizadas, evidenciando respectivos rótulos, siglas correspondentes e descrições sucintas (Seção 2 e Apêndice B, Quadro B. 12).

A última etapa apresenta uma síntese das 28 variáveis utilizadas para carga da ferramenta (Seção 2 e Apêndice B, Quadro B. 5, Quadro B. 14 e Quadro B. 15). Os itens descritos nesse quadro incluem: identificadores das variáveis, rótulos (nomenclatura utilizada), categoria da variável (por exemplo, financeira, de recursos humanos, de produção) e setor da TH, que identifica a qual setor da TH a variável é vinculada. Essas informações ajudam a entender melhor como as variáveis foram definidas e utilizadas no estudo.

### 3.5 Modelo de BI

Conforme ilustrado na Figura 1, uma aplicação de BI pode trazer diversos resultados. Optou-se por aplicar as tecnologias de DW e *dashboard* para criar uma ferramenta *online*, com vistas a atender ao objetivo deste trabalho e, paralelamente, realizar um *data mining* dos dados. Apesar de serem mais restritas que opções OLAP, o *dashboard* permite a visualização de dados e o *data mining* a compreensão, comportamento e identificação de padrões significativos dos dados, trazendo informações úteis a gestores, pesquisadores e comunidade em geral.

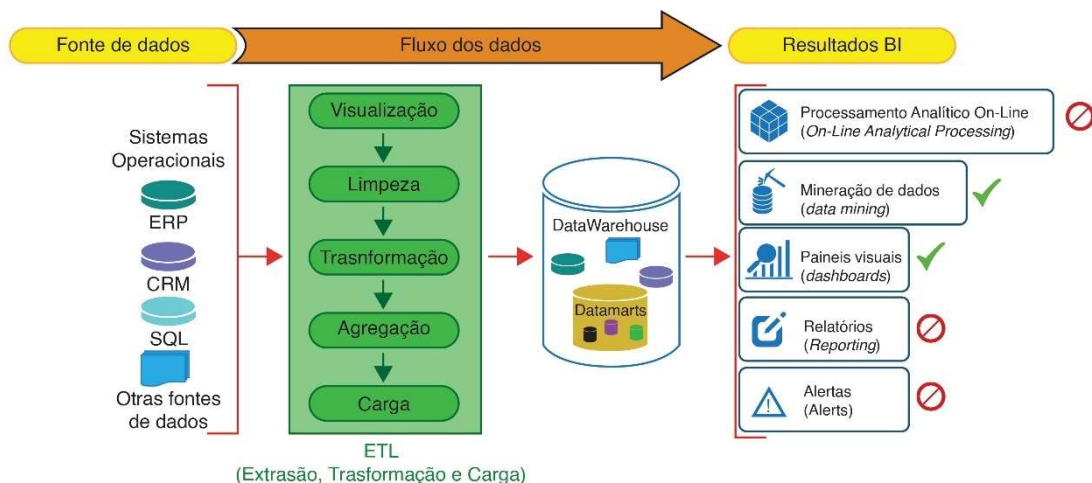


Figura 1: Representação esquemática da técnica clássica de BI a partir de fontes de dados. Adaptado das referências Krmac (2011) e Fana *et al.* (2021), os *checks* positivos identificam os resultados de BI adotados neste trabalho.

A Figura 2 representa esquematicamente a arquitetura BI utilizada neste estudo.

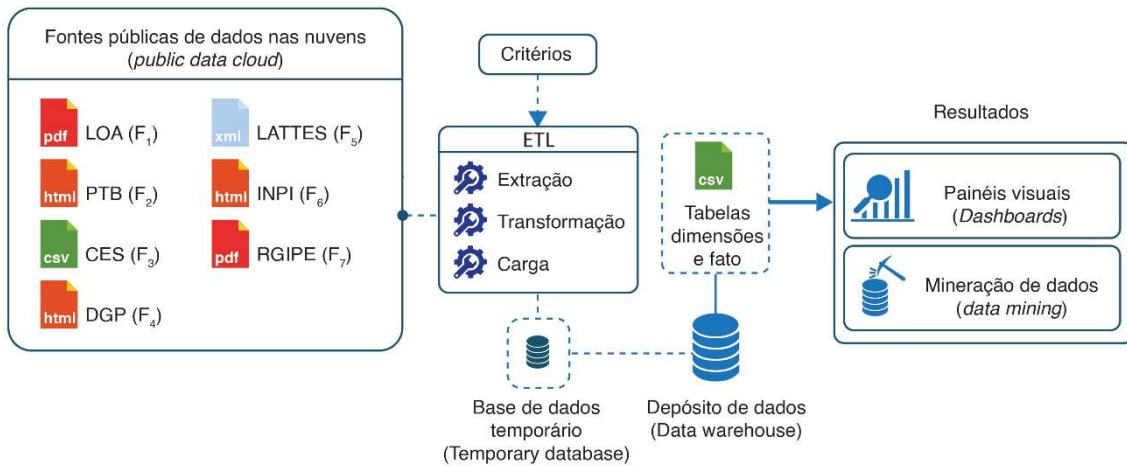


Figura 2: Representação esquemática do modelo de BI utilizado neste trabalho, a partir de dados públicos, adaptada de Krmac (2011) e Fana *et al.* (2021).

Para este trabalho, a técnica de BI (Figura 2) adotada foi dividida em 7 etapas, detalhadas a seguir:

- i. Identificação das fontes de dados abertas do Governo Federal que armazenam informações sobre o ecossistema das ICTs;
- ii. Definição dos critérios de inclusão e exclusão dos dados a serem extraídos;
- iii. Aplicação do procedimento ETL, particularmente as etapas de extração e transformação;
- iv. Elaboração do Modelo de Entidade de Relacionamento temporário;
- v. Armazenamento no Banco de Dados temporário o resultado das etapas anteriores.
- vi. Carregamento dos dados no DW, utilizando o esquema misto;
- vii. Apresentação da criação do *dashboard* desenvolvido;
- viii. Mineração dos dados para a identificação de padrões significativos considerando as características geoespaciais e temporais dos dados analisados.

Observa-se que devido à complexidade do processo, a limitação de tempo e recursos humanos e financeiros disponíveis o procedimento OLAP foi suprimido deste MVP, podendo ser incorporando em versão futura.

### 3.5.1 Fonte de Dados

Várias fontes de dados abertos fornecidas pelo Governo Federal, identificadas durante a análise do referencial teórico (Seção 2 e Apêndice B, Quadro B. 13) foram consideradas. Especificamente, a Lei Orçamentária Anual, o Portal da Transparência e o Portal do Censo Superior da Educação foram adicionadas por serem consideradas significativas, dada a pertinência das informações que fornecem acerca do ecossistema das ICTs. Como critério de identificação dessas fontes, considerou-se a pertinência e relevância dos das variáveis disponíveis em relação aos objetivos de pesquisa e às necessidades de dados para análise. Os dados, coletados dessas fontes apresentavam formatos diversos como *html*, *csv*, *xls*, *xml*, *sql*, *txt* e *pdf*.

Foram identificadas 43 variáveis, frequentemente utilizadas em estudos anteriores (Apêndice B, Quadro B. 5), que foram consideradas como referência para a seleção das 28 variáveis adotadas (Apêndice, Quadro B. 14). Entretanto, devido à falta de padronização e transparência dos dados, somente 12 dessas últimas foram consideradas como compatíveis, após análise.

O procedimento de seleção das 12 variáveis seguiu as seguintes etapas:

- **Revisão da Literatura e Identificação Preliminar:** Inicialmente, foi realizada uma revisão abrangente da literatura e estudos anteriores, listados no Apêndice B, Quadro B.5, que identificou 43 variáveis frequentemente utilizadas em pesquisas relacionadas aos canais de PD&S das ICTs.
- **Seleção Baseada em Critérios de Relevância:** Dessas 43 variáveis, 28 foram pré-selecionadas com base em sua relevância e alinhamento com os objetivos da pesquisa atual. Esse processo de seleção foi guiado por critérios definidos previamente, como a pertinência das variáveis aos temas de inovação, desenvolvimento tecnológico e levantamento de requisitos dos usuários consultados.
- **Análise da Disponibilidade e Compatibilidade dos Dados:** Após a seleção inicial, as 28 variáveis foram submetidas a uma análise detalhada quanto à disponibilidade e compatibilidade dos dados. Essa etapa envolveu a verificação da presença de informações correspondentes nas fontes de dados abertos fornecidas pelo Governo Federal, considerando a padronização, acessibilidade e qualidade dos dados disponíveis.

- **Constatação da Falta de Padronização e Transparência:** Durante a análise, foi constatada a falta de padronização e transparência de alguns dados disponíveis. Isso implicou diferenças nos formatos de dados, unidades de medida, definições de variáveis e a ausência de informações cruciais para identificar as variáveis selecionadas, tornando desafiador o uso consistente dessas informações para análise.
- **Seleção Final de Variáveis Compatíveis:** Devido aos desafios identificados com a padronização e a transparência, somente 12 das 28 variáveis selecionadas foram consideradas como compatíveis para uso na pesquisa. Essas variáveis foram as que apresentaram informações disponíveis de maneira clara, estruturada e acessível nas fontes de dados governamentais, permitindo uma análise consistente e alinhada com os objetivos do estudo.

Esse procedimento assegurou que as 12 variáveis finais fossem não apenas relevantes ao escopo do estudo, mas também apoiadas por dados confiáveis e consistentes, permitindo análises precisas e significativas.

Posteriormente, 7 variáveis adicionais foram integradas ao conjunto, pois elas não foram identificadas na revisão teórica, mas foram citadas durante o levantamento dos requisitos, das quais 3 são discretas, 1 contínua e 3 nominais, resultando em um total de 19. Dentre as 11 discretas, foram derivadas mais 9, com o objetivo de fornecer informações sobre suas interações TH. Tais adições englobam variáveis como *programas de computadores associados aos SPP* ( $X_{24}$ ) e a *cessão de infraestrutura solicitada pelo SPP* ( $X_{17}$ ). Portanto, este estudo considera um total de 28 variáveis para desenvolvimento da solução tecnológica em BI, proposta no objetivo geral.

Os dados correspondentes às 28 variáveis foram então extraídos das 7 fontes distintas consideradas (Apêndice B, Quadro B. 13). A fim de estabelecer a compatibilidade entre essas fontes de dados, diversos procedimentos foram postos em prática. Um dos procedimentos realizados diz respeito ao INPI e requereu a execução de uma etapa unificadora. Durante o período entre a coleta e o carregamento dos dados, uma consulta adicional à base de dados da Receita Federal foi necessária para determinar o Cadastro Nacional de Pessoa Jurídica (CNPJ) de cada ICT, abrangendo tanto a matriz quanto as filiais. O CNPJ, em razão de sua precisão e registro no INPI, foi empregado como chave primária, conectando, deste modo, as informações coletadas entre os registros contidos no INPI, cuja titularidade pertenciam a cada ICT da RFEPCT.

A efetivação destes procedimentos assegurou a coesão entre os diferentes conjuntos de dados, viabilizando assim uma análise mais integrada e coerente dos dados.

### **3.5.2 Critérios para Inclusão e Exclusão dos Dados**

Na segunda etapa, complementada por um estágio adicional, definiram-se os critérios para inclusão e exclusão de dados oriundos das 7 fontes escolhidas, a partir do levantamento de requisito realizado (Subseção 3.2). O critério de inclusão visou assegurar que os dados escolhidos contribuam de forma significativa para três aspectos fundamentais: i) a geração de conhecimento, ii) o aperfeiçoamento do processo de tomada de decisão, e iii) a formulação de políticas públicas institucionais de PD&S. O objetivo estratégico desse processo de seleção foi assegurar a relevância e a aplicabilidade dos dados coletados.

Paralelamente, o critério de exclusão de dados buscou garantir que, mesmo quando estes estivessem associados ao contexto das ICTs, seriam excluídos, se não apresentassem um padrão claro que assegurasse sua confidencialidade, integridade e disponibilidade.

Projetos<sup>13</sup> de extensão tecnológica ou tecnológicos de extensão, por exemplo, foram excluídos devido às diversas interpretações atribuídas por gestores e à sua pouca presença em estudos, outros pontos a se considerar é a confidencialidade, integridade e disponibilidade desses dados nas fontes utilizadas.

### **3.5.3 Extração, Transformação e Carga dos dados selecionados para o projeto**

O processo de ETL, essencial na elaboração de uma eficaz solução BI, compreende três etapas distintas: extração, transformação e carga (Apêndice E, Apêndice B, Figura B. 3), no qual, parte dele foi omitida pelos motivos previamente mencionados (Seção 3).

Na etapa inicial de extração, visa-se coletar dados de múltiplas fontes, como bancos de dados, planilhas e sistemas externos. Essa fase é crucial para reunir

---

<sup>13</sup> Destaca-se que, durante o período de colaboração com a gestão da PRPI, em uma determinada consulta do FORPOG, para atualização dos dados da Rede, surgiu a necessidade de entendimento sobre qual classificação seria adotada para esse item e outros itens, tais como o tipo de pesquisa e de patente. Havia, pois, uma generalização daquilo que deveria ser classificado por segmento, por falta de um entendimento conceitual, ou seja, apenas baseado na subjetividade dos gestores.

informações diversificadas e relevantes ao processo analítico, a qual demandou cerca de 80% de todo o trabalho.

Subsequentemente, na fase de transformação, ocorre a limpeza, padronização e integração dos dados coletados. Essa etapa assegura a consistência e a qualidade das informações, evitando inconsistências que possam prejudicar as análises futuras.

Finalmente, na etapa de carga, objetivou-se carregar os dados já processados e uniformizados em um conjunto de repositório de dados específico (*Data Mart*), visando facilitar o acesso e a análise dessas informações, garantindo sua integração e otimização no armazenamento centralizado do *Data Warehouse*. Concluído esse processo, os dados estavam prontos para serem analisados e aplicados na tomada de decisões.

Destaque-se que a qualidade do resultado do processo de ETL depende da qualidade dos dados originais e da competência dos profissionais envolvidos. Portanto, foi fundamental dedicar tempo e atenção adequados para garantir que os dados estejam bem-preparados para análise e tomada de decisões.

#### **3.5.4 Modelo de Entidade-Relacionamento dos Dados**

Para a criação do MER, utilizou-se o SGDB, denominado de MySQL, na versão 5.6.36-82.0-log. A interface visual foi proporcionada pelo MySQL *Workbench* na versão 8.0.28 e a codificação foi realizada, utilizando-se a linguagem SQL. A adoção do Modelo de Entidade-Relacionamento e do MySQL foram fundamentais para a integração e consolidação de dados provenientes de diferentes fontes integradas neste trabalho. Parte do processo foi omitida pelos motivos previamente mencionados (Seção 3).

A construção do Modelo de Entidade-Relacionamento permitiu uma melhor compreensão das inter-relações dos dados entre as tabelas, que foram organizadas, a partir de várias fontes de dados integradas (Apêndice B, Quadro B. 13 e Figura B. 1), após um conjunto de cargas entre as etapas 3 e 4, um exemplo ilustrativo pode ser acessado no Apêndice B, Código B. 1.

#### **3.5.5 Implementação Híbrida de Data Warehouse: integrando Modelos Estrela e Floco de Neve para Otimização de Desempenho e Análise**

A construção DW baseado em um modelo misto de Modelo Estrela e Floco de Neve requereu um planejamento e execução meticulosa que envolveu várias etapas



fundamentais. Parte do processo foi omitida pelos motivos previamente mencionados (Seção 3).

O primeiro passo envolveu um levantamento das necessidades e requisitos dos usuários (Subseção 3.2), a fim de compreender quais dados deveriam ser armazenados no DW e como seriam utilizados, na versão MVP desenvolvida nesse trabalho.

Em seguida, os dados relevantes foram identificados, coletados, validados e carregados, a partir das fontes de dados originais (Apêndice B, Quadro B. 13). Uma vez que os dados foram coletados, eles foram submetidos ao processo de ETL (Subseção 3.5.2), no qual os dados foram extraídos das fontes/sistemas definidas, e passaram por processos de limpeza, padronização e integração. Em sequência ocorreu o design da estrutura do DW, incluindo a criação de Tabelas Dimensões e Fatos que facilitaram as consultas e análises (Apêndice D e Apêndice B, Quadro B. 6). Um modelo híbrido de Modelo Estrela e Floco de Neve foi implementado, subsequentemente.

O modelo Estrela, com a tabela de fatos no centro, rodeada por tabelas de dimensão, foi utilizado para atender às consultas que precisam de respostas rápidas, por exemplo o Módulo **Amplo** (Apêndice B, Figura B. 15).

O modelo Floco de Neve, que é uma versão normalizada do modelo Estrela, com tabelas de dimensão adicionais, foi aplicada para consultas que requeriam um nível mais detalhado de análise, por exemplo os módulos **Detalhado** (Apêndice B, Figura B. 16) e **Comparativo** (Figura B. 17).

Posteriormente, o DW foi populado com os dados históricos e ficou pronto para fornecer suporte à análise de negócios. Por fim, a manutenção e o monitoramento contínuo do DW são fundamentais para garantir a qualidade dos dados e a eficiência do sistema, acompanhando a evolução das necessidades do negócio ao longo do tempo.

Caracterizado esse produto como um MVP, o processo de atualização automática foi suspenso, contudo, pode ser reativado e aprimorado, caso utilizado efetivamente por uma ICTs. Assim, a construção do DW envolveu uma combinação desses dois modelos, atendendo tanto às necessidades de desempenho quanto de análise detalhada.

### 3.5.6 Tabelas Dimensão e Fato

A Figura 3 exemplifica como as Tabelas Dimensões e Fato são apresentadas na interface (Apêndice D). As Tabelas Dimensões são representadas visualmente pelos filtros básicos disponíveis destacados pela área contida na área delimitada pela linha

tracejada. A Tabela Fato é representada visualmente pela área delimitada pela linha traço dois pontos. As áreas listam os resultados em formato de tabela para exportação, a depender do perfil do usuário, a partir dos filtros realizados por ele. Para uma melhor experiência de usuário, é possível acessar a ferramenta *online*<sup>14</sup>.

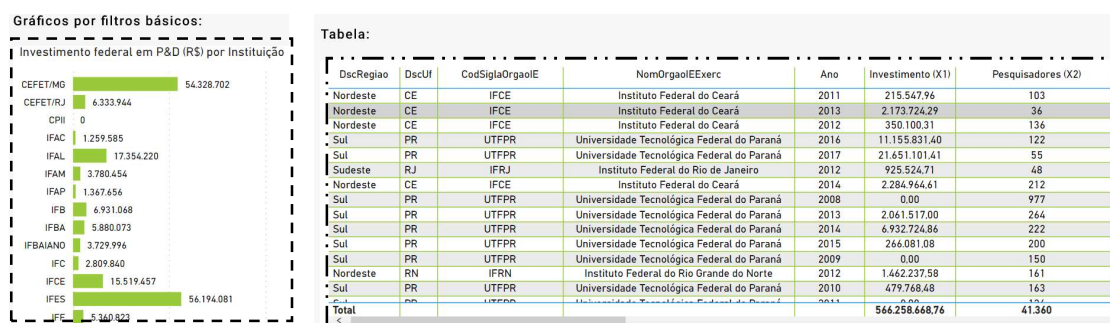


Figura 3: Ilustração das tabelas dimensões (representada pelas áreas de linha tracejada, filtros) e fato (representada pela área de linha traço dois pontos, métricas) para o período de 2008 a 2018 da RFEPCT, disponibilizado no MVP.

A Figura 3 e a Tabela B. 1 (Apêndice B) ilustram os filtros disponíveis (área tracejada) para múltiplas seleções e para as métricas (área traço dois pontos). Os atributos são divididos em 4 tipos: hierárquicos, por exemplo, estão representados nas colunas 1 e 2; descritivos são exemplificados nas colunas 3 e 4; temporais evidenciados na coluna 5 e, finalmente, os atributos de medida que se encontram a partir da coluna 6 até a 24. Toda essa informação, extraída das Tabelas Dimensões e Fatos, facilita uma análise organizada dos dados.

As colunas 1 e 2, da Figura 3, representam os atributos referentes à localização geoespacial das ICTs, no território nacional. As colunas 3 e 4, da Figura 3, representam os atributos referentes às siglas e descrição das ICTs analisadas. A coluna 5, da Figura 3, representa o atributo referente ao ano que os dados correspondem.

Por fim, as colunas de 6 a 24, Tabela B. 2 (Apêndice B), representam os atributos referentes aos indicadores:  $X_1$  (coluna 6), que expõem os recursos públicos federais destinados ao fomento dos canais P&D;  $X_2$  (coluna 7), que apresenta a quantidade de pesquisadores;  $X_3$  a  $X_9$  (colunas entre 8 e 14), que possuem registros sobre aos canais de pesquisa (gerais ou vinculados ao atendimento das demandas dos SPP);  $X_{10}$  a  $X_{17}$  (colunas entre 15 a 22), que apresentam registros sobre os serviços acadêmicos (gerais ou vinculados ao atendimento das demandas dos SPP);  $X_{20}$  a  $X_{24}$  (colunas entre 23 a 29),

<sup>14</sup> <https://prequest.websiteseuro.com/bi>

que exibem dados sobre os resultados de desenvolvimento tecnológico (acadêmicos ou vinculados ao SPP). Tais atributos podem ser aplicados em qualquer ecossistema de ICTs.

### 3.5.7 Indicador-chave de Desempenho (KPI)

A construção dos KPI para este DW baseou-se principalmente nas variáveis-chave indicadas e nos requisitos do usuário. Para cada variável, foi realizada uma série de etapas para garantir a qualidade e a relevância dos dados. Inicialmente, foi necessária uma compreensão clara da definição e do significado de cada variável. Por exemplo, "Região", "UF" e "Órgão" são relativamente diretos, referindo-se a divisões geográficas e a entidades governamentais específicas. Variáveis como "Investimento", "Pesquisadores", "Orientação de Pesquisa", "Projeto de Pesquisa", entre outras, foram definidas em colaboração com os principais *stakeholders*, como o manual<sup>15</sup> de utilização e preenchimento do Currículo Lattes e o módulo instituição de educação superior<sup>16</sup> para garantir que refletissem as necessidades de negócios. A síntese das descrições para cada uma das 28 variáveis-chave pode ser consultada no (Apêndice B, Quadro B. 13).

Uma vez definidas as variáveis/métricas, foram identificadas as fontes de dados relevantes para cada uma delas (Apêndice B, Quadro B. 13). Isso envolveu a integração com sistemas existentes, conforme mencionado no requisito dos usuários (Subseção 3.2: requisito i), garantindo que os dados pudessem ser coletados de maneira eficiente e confiável.

Em relação aos requisitos dos usuários, a construção dos KPI buscou oferecer as ferramentas necessárias à segmentação e análise dos aspectos relevantes aos canais de PD&S acessados. Para demonstração da potencialidade e flexibilidade do MVP proposto, outros exemplos de KPI foram adicionados. Embora relevante, a criação de relatórios não foi inclusa nesta versão do MVP, devido à restrição de tempo e recursos financeiros. Tal inclusão está prevista para etapas futuras do projeto.

---

<sup>15</sup> Disponível em: [http://ajuda.cnpq.br/index.php/P%C3%A1gina\\_principal](http://ajuda.cnpq.br/index.php/P%C3%A1gina_principal) , acesso em 03 fev. 2017.

<sup>16</sup> Disponível em: [https://download.inep.gov.br/educacao\\_superior/censo\\_superior/manuais\\_e\\_instrucoes/manual\\_de\\_IES\\_2022-versao\\_preliminar.pdf](https://download.inep.gov.br/educacao_superior/censo_superior/manuais_e_instrucoes/manual_de_IES_2022-versao_preliminar.pdf), acesso em 09 jan. 2023.

### 3.5.8 *Dashboard*

Definidos os KPI, a etapa subsequente foi a elaboração dos *dashboards* interativos, que facilitam a visualização, exploração e análise de dados. Essas ferramentas apresentam informações de forma dinâmica, geradas a partir de perguntas formuladas pelos usuários e contextualizadas por eles mesmos, comumente exibidas em formato gráfico.

Os *dashboards* possibilitam a avaliação sob múltiplas perspectivas, tais como a performance dos pesquisadores (Subseção 3.2: requisito iii), o desempenho das PD&S em cooperação com SPP (Subseção 3.2: requisito v), e o investimento em PD&S (Subseção 3.2: requisito vi), tornam-se ferramentas valiosas na interpretação de informações, por meio da interação do usuário.

Portanto, os *dashboards* podem ser considerados como um componente vital do sistema *web* baseado em BI, permitindo a exploração de informações sob diversas perspectivas e promovendo uma compreensão intuitiva e abrangente das métricas de desempenho.

No contexto apresentado, o sistema foi estrategicamente planejado em 3 módulos – **Amplio**, **Detalhado** e **Comparativo** – com o intuito de permitir consultas personalizadas, por meio da filtragem de dados por variáveis específicas (Apêndice B, Figura B. 15-Figura B. 17). Esse arranjo facilita a aquisição rápida e intuitiva de *insights* relevantes, através de um *UX design* funcional, cumprindo assim o requisito vii (Subseção 3.2). Embora relevante, a estética visual está simplificada nesta versão do MVP devido à restrição de recursos e de tempo. Tal melhoria está prevista para etapas futuras do projeto.

O *UX design* do *dashboard*, proposto neste trabalho, foi planejado de modo a incorporar filtros adicionais e setores que ampliam as possibilidades de interação e de extração de análises potenciais relacionadas a métricas e/ou KPI de forma dinâmica. Portanto, ao projetá-lo, buscou-se favorecer a realização de análises mais abrangentes e interativas, elevando assim as oportunidades para extrair *insights* de métricas e KPIs de maneira ágil e eficaz.

Para o desenvolvimento da solução proposta, optou-se pelo formato retrato com o objetivo de garantir uma visualização ideal das informações em dispositivos móveis, como *smartphones* e *tablets*. Essa escolha proporciona um acesso mais flexível e conveniente aos dados analisados. No entanto, salienta-se que o *UX design* pode ser

adaptado futuramente para outros formatos, como o modo paisagem, além de incluir novas variáveis, métricas e KPIs, de acordo com as demandas dos usuários.

### 3.5.9 *Data Mining*

O processo de *data mining* (Apêndice C) busca explorar e analisar grandes volumes de dados a fim de identificar padrões e regras significativas. Para isso, optou-se pelo uso de duas técnicas de interdependência multivariada, amplamente reconhecidas em diversos campos científicos. Os resultados estatísticos das análises foram relatados de acordo com as diretrizes da *American Psychological Association* (APA), fornecendo informações precisas para uma avaliação rigorosa.

A análise de agrupamento se divide em duas possibilidades: Análise de *Clusters* Hierárquica (em inglês, HC) e Não Hierárquica (em inglês, NHC). Na primeira etapa, a análise de HC buscou identificar a taxonomia, e, a análise de NHC, as segmentações da RFEPCT, com base nas suas (dis)similaridades. Para tanto, foram consideradas as variáveis/métricas selecionadas neste estudo (Apêndice B, Quadro B. 14), divididas em (23) vinte e três discretas e (1) uma contínua, correspondendo à produção dos canais de PD&S das 42 ICTs, após a primeira década de atividades, disponíveis nas fontes de dados consultadas (Apêndice B, Quadro B. 13).

A segunda foi a análise fatorial exploratória (em inglês, EFA) por componentes principais (em inglês, PCA), cujo objetivo foi identificar o conjunto de variáveis latentes das variáveis acima apresentadas. Na sequência foi proposto um indicador de desenvolvimento PD&S, aplicado as ICTs. Ele considerou os componentes principais identificados de maneira multidimensional e multivariada pelo procedimento PCA, a partir dos dados de PD&S disponíveis no MVP (Subseção 3.5.6).

A medida de distância considerada para classificar uma possível ICT como *outliers* multivariada foi a distância de *Mahalanobis*, a qual é utilizada como valor padrão do algoritmo BACON, cujo valor de corte adotado foi um *alpha* de 0,15 (Billor *et al.*, 2000; Fávero *et al.*, 2020, p.375). Para identificar possíveis *outliers* univariados, foi empregada a técnicas estatísticas de ANOVA *one way*, que é utilizada para verificar a existência de diferenças significativas entre grupos, baseando-se em um valor de corte para o *p-valor* estabelecido em  $> 0,05$  (Fávero *et al.*, 2020, p. 338).

### 3.5.9.1 *Análise de Agrupamento*

A análise de *clusters* é um método popular na análise multivariada, exploratória, que visa identificar grupos de dados com características semelhantes (Apêndice C, Subseção C.1). Eles se dividem em HC e NHC. No trabalho em questão, essa abordagem foi empregada para interpretar o conjunto de dados gerados com base na seleção padrão, cuja representação se encontra na última tabela disponível no módulo **Amplo** do MVP (Apêndice B, Figura B. 15). Salienta-se que a inserção de novos dados gera novos e outros resultados.

Essa tabela foi exportada para a linguagem *R*, cumprindo assim um dos objetivos complementares, a saber: utilizar a análise de agrupamentos para identificar a taxonomia e as características da RFEPCT, com base nas 42 ICTs, após a primeira década de atividades, a partir dos fatores identificados (Subseção 3.5.9.1, Item 7).

Para alcançá-los, foram realizados os seguintes procedimentos:

1. **Análise de Agrupamento Hierárquico (em inglês, *Hierarchical Clustering - HC*):**
  - a. selecionar as variáveis (Subseção 3.5.9.1, Item 7);
  - b. calcular matriz de distância euclidiana;
  - c. aplicar o método de ligação *Ward's*;
  - d. construir o dendrograma;
  - e. definir o ponto de corte do dendrograma;
  - f. determinar o número de *Clusters* inicial;
  - g. aplicar o teste estatístico de ANOVA;
  - h. determinar o número de HC final.
  
2. **Análise de Cluster Não Hierárquico (NHC *K*-médias):**
  - a. selecionar as variáveis (Subseção 3.5.9.1, Item 7);
  - b. calcular matriz de distância euclidiana;
  - c. definir o valor de *K* inicial (Subseção 3.5.9.2, Item 1; subitem f);
  - d. selecionar a melhor centroide;
  - e. aplicar o teste estatístico de ANOVA inicial;
  - f. determinar o número de NHC final.

Para as análises HC e HCA, decidiu-se abordar os resultados sob um critério geral e duas abordagens distintas, a serem discutidas a seguir. Essa decisão foi tomada com o propósito de ampliar e fortalecer a compreensão dos dados.

O primeiro critério adotado foi que, para ser considerado um grupo na análise HC ou HCA, deveria conter no mínimo duas observações, o equivalente a duas IEPTC. Uma vez que esse critério não fosse atingido, em qualquer ano ou intervalo temporal, grupos individuais seriam excluídos da análise, por representarem possíveis *outliers* tão significativos a ponto de não permitirem a criação de, pelo menos, dois grupos com um mínimo de duas observações.

As duas abordagens referem-se à escolha de incluir ou excluir variáveis classificadas como "não contributivas" na análise. Essa prerrogativa é do pesquisador, pautada em sua compreensão analítica e técnica do "modelo de negócio" (Fávero *et al.*, 2020; Field, 2018; Hair *et al.*, 2019). Com base nessa premissa, optou-se por analisar as seguintes abordagens:

- Abordagem 1: considerou-se as 24 variáveis no estudo, ainda que possuíssem ( $\mu$  e  $\sigma = 0$ ).
- Abordagem 2: considerou-se apenas as variáveis cujas ( $\mu$  e  $\sigma \neq 0$ ) e aquelas constatadas como "contributivas", após ANOVA.

A escolha das duas abordagens possibilitou uma avaliação completa dos resultados, levando-se em consideração tanto as influências estatisticamente significativas quanto a exclusão daquelas consideradas não significativas. Isso proporcionou *insights* mais abrangentes e robustos sobre a RFEPCT.

Durante o processo de análise hierárquica ou não hierárquica, ao deparar-se com uma ocorrência de singularidade matricial que impeça a obtenção de uma solução válida, optou-se por excluí-la para permitir a continuidade da análise (Fávero *et al.*, 2020; Hair *et al.*, 2019). Essa exclusão é justificada pela necessidade de obter uma solução estável e confiável para a análise hierárquica e assegura que os resultados obtidos são derivados de uma matriz que é invertida e, portanto, válida para análises futuras.

Esse fenômeno é caracterizado por uma matriz que não possui sua forma inversa, muitas vezes levando a erros de computação e resultados não confiáveis. Ele pode surgir devido a problemas como multicolinearidade, em que uma variável pode ser prevista perfeitamente a partir de outras variáveis ou a partir da presença de *outliers* extremos.

É pertinente destacar que tal exclusão, embora necessária para a progressão da análise, pode influenciar os resultados e a interpretação dos dados. Assim, recomenda-se cautela durante a interpretação dos resultados e, se possível, a realização de análises adicionais para validar as conclusões derivadas da matriz modificada.

Nos procedimentos de análise HC e NHC da Abordagem 2, foi utilizado o Índice Ajustado de Rand (em inglês, ARI) para avaliar a similaridade entre os agrupamentos formados em cada método. O ARI é uma métrica que quantifica a concordância entre dois agrupamentos, variando de -1 a 1, em que um (1) indica que os agrupamentos são idênticos e o zero (0) indica que o agrupamento é resultado de uma atribuição aleatória de *clusters* (Hubert *et al.*, 1985).

### 3.5.9.2 *Análise Fatorial Exploratória por Componentes Principais*

A EFA por PCA tem o objetivo de reduzir a dimensão de um conjunto de dados, ao mesmo tempo preservando a máxima variância possível (Apêndice C, Subseção C.2). No trabalho em questão, essa abordagem foi empregada para interpretar o conjunto de dados gerados com base na seleção padrão, cuja representação se encontra na última tabela disponível no módulo **Amplo** do MVP (Apêndice B, Figura B. 15). Essa tabela foi exportada para a linguagem *R*, cumprindo assim um dos objetivos complementares, a saber:

- i) identificar as correlações entre as variáveis originais para criação de componentes principais que representem suas possíveis combinações, buscando reduzir a dimensionalidade da matriz dos dados a serem analisados;
- ii) acompanhar a RFEPCT através de suas 42 ICTs, após a primeira década de atividades, utilizando o indicador de PD&S, a partir dos componentes principais identificados.

Para alcançá-los foram realizados os seguintes procedimentos:

1. selecionar as variáveis (Subseção 3.4);
2. preparar os dados (Subseção 3.5.3);
3. aplicar o teste estatístico de esfericidade de Bartlett;
4. encontrar Autovalores e Autovetores;
5. ordenar Autovetores por Autovalores Decrescentes;



6. extrair os Componentes Principais;
7. determinar o número de Componentes;
8. efetuar a Análise Fatorial;
9. interpretar os Componentes;
10. elaborar o indicador de desenvolvimento PD&S.

As características de média, mediana e desvio padrão das variáveis foram documentadas, assim como os componentes principais derivados do PCA, juntamente com as respectivas variâncias explicadas. Esses resultados quantitativos enriqueceram a compreensão do intervalo temporal entre 2008 e 2018 e constituíram uma base sólida para interpretações posteriores.

### **3.6 Validação da Solução**

Para validar o MVP proposto, seguiram-se as etapas metodológicas específicas: revisão, teste e feedback de especialistas (Apêndice B, Quadro B. 3 e Apêndice I), relacionadas com as teorias da Triple Helix de inovação, teoria da Análise de Negócios e teoria da Visualização de Dados.

Inicialmente, revisou-se minuciosamente as especificações do sistema (Subseção 3.2), garantindo a correta definição e alinhamento dos requisitos às necessidades do negócio. Essa revisão fundamenta-se na teoria da Análise de Negócios, que enfatiza a importância de um alinhamento preciso entre as soluções tecnológicas e os objetivos empresariais.

Na sequência, foram aplicados diversos testes – unidade, integração, aceitação do usuário, validação de análises de dados e outros – com o propósito de assegurar a funcionalidade, confiabilidade e utilidade do sistema no processo decisório. Tais testes têm sua importância sublinhada pela teoria da Visualização de Dados, pois garantem que o sistema opera conforme o previsto e fornece resultados confiáveis para tomada de decisões e geração de *insights*.

Por fim, fundamentando-se na teoria da Triple Helix de inovação, que destaca a colaboração entre ICTs, SPP e GOV, identificou-se a relevância de validar a usabilidade, crucial para a eficácia de sistemas baseados em BI. Portanto, testes de usabilidade e observações diretas foram realizados, assegurando facilidade e eficiência no uso do MVP.

Os *feedbacks* recolhidos tiveram o intuito de validar ou refutar a questão de pesquisa, o objetivo geral, e a hipótese apresentada, permitindo refinar o sistema em sua versão MVP. Ainda que esta versão não incorpore todos os *feedbacks* sugeridos, planeja-se a inclusão destas contribuições em futuras melhorias.

## 4 Resultados

Esta seção apresenta os resultados deste trabalho. Na primeira subseção, é apresentada a solução tecnológica proposta (MVP) tendo como referência os requisitos dos usuários, os pressupostos conceituais objetivos relacionados às teorias da TH, da Análise de Negócio e da Visualização de Dados e a aplicação do *UX design* planejado.

Na segunda subseção, são apresentados os principais resultados obtidos a partir da comparação da RFEPCT com redes similares em outros países.

Na terceira subseção, são expostos os principais resultados relacionados aos canais de PD&S da RFEPCT, desenvolvidos entre 2008 e 2018.

Na quarta, é apresentada uma síntese da EDA relativa às 24 variáveis quantificáveis, relacionadas aos canais de PD&S, que serviram como referência para o desenvolvimento deste MVP.

Na quinta subseção, são apresentados alguns exemplos que simulam o uso do MVP para a análise de dados e a tomada de decisões por um gestor, demonstrando a capacidade da plataforma de atender, em tempo hábil, a elaboração dos relatórios solicitados pelos órgãos de controle governamentais, por exemplo.

Na sexta subseção, são apresentados dois exemplos de *data mining* (Análise de Agrupamento e Fatorial por Componentes Principais), a partir das tabelas geradas mediante os filtros desejados pelo usuário no MVP.

Na última subseção, são apresentados os resultados dos especialistas, os quais permitiram validar e responder à hipótese e a questão de pesquisa apresentada nessa tese.

O MVP proposto oferece informações detalhadas sobre 24 variáveis quantitativas relevantes, relacionadas aos canais de PD&S. Facilita análises comparativas, de correlações entre as métricas disponíveis, tornando-se uma ferramenta valiosa para a tomada de decisões informadas e para a identificação de tendências e padrões relevantes.

Os benefícios entregues por ele englobam a gestão do conhecimento produzido, o estímulo à produção científica e tecnológica, o direcionamento de pesquisas com potencial de impacto socioeconômico e o fortalecimento das relações entre os atores da TH, contribuindo assim para o desenvolvimento regional e nacional.

Especialistas identificam o MVP proposto como uma ferramenta tecnológica essencial e inovadora para enfrentar os desafios do cenário brasileiro de CT&I e PD&I nas ICTs. Possui uma interface intuitiva que favorece a visualização e interpretação dos dados, proporcionando *insights* valiosos que fundamentam decisões estratégicas.

#### **4.1 Comparando Alguns Resultados da Educação Profissional e Tecnológica: um Estudo de Caso da RFEPCT Brasileira e Redes em Países como Alemanha, Canadá, Chile, China e Finlândia**

A comparação da RFEPCT com as redes de educação profissional e tecnológica de países como Alemanha, Canadá (Montreal), Chile, China e Finlândia permite entender as melhores práticas e estratégias que promovem a educação profissional, considerando os resultados notáveis que essas nações obtiveram em P&D (Bank's, 2020). Em relação como a organização da educação profissional ocorre em cada país percebe-se que possui características distintas.

Identificou-se que na Alemanha prevalece o modelo dual, combinando aprendizado teórico e treinamento prático (Brasil, 2016b, p. 25-46); em Montreal, Canadá, a oferta se dá por centros de formação e escolas técnicas (Brasil, 2016b, p. 76-86); no Chile, em escolas técnico-profissionais e institutos técnicos (Brasil, 2016b, p. 137-164); na China, em escolas secundárias profissionais e outros centros (Brasil, 2016b, p. 153-164); e na Finlândia, por instituições vocacionais e programas locais (Brasil, 2016b, p. 275-282).

Ainda em relação a organização, destaca-se que as semelhanças entre os países implicam a identificação da presença de instituições especializadas e colaborações entre os atores propostos pela teoria da TH. Reconhecer tais diferenças exige atenção ao modelo de educação de cada país, como o modelo dual alemão (Brasil, 2016b, p. 25-46) ou os centros de formação profissional em Montreal (Brasil, 2016b, p. 76-86).

Em relação como a regulação da educação profissional ocorre em cada país também varia. Na Alemanha, é centralizada, enquanto em Montreal, é provincial. No Chile, verifica-se uma abordagem descentralizada, e na China, uma centralizada. A Finlândia adota uma regulação compartilhada.

Em relação como o financiamento da educação profissional ocorre em cada país também revela abordagens distintas. Países como Brasil, Alemanha e Finlândia optam por um forte investimento público, enquanto Chile e China mostram maior diversificação, incluindo participação privada.

Ao observar a abrangência da educação profissional, verifica-se que países como Alemanha e Finlândia abrangem diversos setores e níveis, enquanto Chile e China concentram-se em focos específicos.

Ao empregar a teoria TH para ajudar no entendimento do fenômeno estudado, identificou-se que a organização da RFEPCCT é análoga à dos países estudados (Brasil, 2016b), pois todos aderem à teoria TH (Etzkowitz *et al.*, 1995; 2000). Essa perspectiva permitiu analisar de forma objetiva (dis)similaridades relacionadas à organização, regulação, financiamento e abrangência entre os países.

Os resultados enfatizam a singularidade da RFEPCCT (Brasil, 2008) em termos globais, mostrando sua atuação em mais níveis de formação, em comparação com outros países (Brasil, 2016b). Compreender sua complexidade permite aos gestores propor ações e políticas institucionais que busquem não só atender às demandas dos órgãos públicos, mas também satisfazer todas as demandas da sociedade e fomentar o desenvolvimento socioeconômico.

## **4.2 Resultado do Desenvolvimento do MVP**

Nesta etapa, será fornecida uma visão geral do sistema desenvolvido no *Power BI* para criar um *dashboard*. A solução deve ser capaz de integrar as diferentes fontes de dados identificadas e oferecer uma visualização objetiva e confiável das informações referentes aos canais de PD&S em ecossistema das ICTs.

### **4.2.1 Passo 1: Integração dos Dados**

Para integrar os dados no MVP proposto, foi necessário inicialmente realizar um procedimento de ETL (Subseção 3.3.3). A escolha para armazenamento dos dados recaiu sobre o sistema de gerenciamento de banco de dados MySQL (Subseção 3.5.4). Apesar de sua capacidade, o seu uso em uma solução de DW requer a implementação de um modelo dimensional, essencial para a otimização da análise de dados.

O *software Power BI* foi empregado para a realização da modelagem dimensional, bem como para a segunda etapa do processo de ETL. Este último envolveu a extração e a pré-organização dos dados em um banco de dados temporário, viabilizando uma análise mais eficiente e detalhada no contexto da solução de DW.

As Figura B. 4 a Figura B. 6 (Apêndice B) delineiam os passos para carregar os dados no *Power BI*. Esses passos incluem desde o esquema geral de carregamento até a seleção do banco de dados MySQL, e seu subsequente acesso, para efetuar o carregamento dos dados no *Power BI*.

Após estas etapas, os dados foram então carregados no DW, modelado através do *Power BI*. Foi concebido para fornecer uma visualização objetiva e confiável das informações pertinentes aos canais de PD&S em ICTs (Apêndice B, Figura B. 7 e Figura B. 8).

As Figura B. 7 e Figura B. 8, supracitadas, ilustram a evolução do projeto de DW no *Power BI*. A Figura B. 7 exibe as tabelas originais pré-tratadas, geradas pelo primeiro procedimento de ETL no MySQL, evidenciando a ausência de modelagem dimensional na fase inicial do projeto. Adicionalmente, a Figura B. 8 apresenta as tabelas de dados após o tratamento efetuado no *software* utilizado, durante a segunda etapa de ETL. Destaca as diversas modelagens dimensionais implementadas no DW ao término do projeto, com suas respectivas Tabelas de Fatos e Dimensões.

Esta etapa ressalta a importância da modelagem dimensional em um projeto de DW, que é crucial para permitir uma análise e compreensão dos dados de maneira mais eficaz e intuitiva.

#### **4.2.2 Passo 2: Modelagem dos Dados**

Após a integração dos dados, optou-se pela utilização do *Power BI* para a realização da modelagem dimensional e do processo de ETL. A linguagem *M script*, nativa desse *software*, foi empregada para executar transformações e limpezas de dados. Sublinha-se a importância de criar modelos dimensionais alinhados aos requisitos dos usuários, visando assegurar a efetividade da solução de DW. Essa etapa foi considerada a mais crítica, pois demandou aproximadamente 80% do tempo de desenvolvimento do MVP.

A Figura B. 9 (Apêndice B) exemplifica uma das várias modelagens dimensionais implementadas no DW para o desenvolvimento do projeto no *Power BI*. Por exemplo, essa implementação, crucial para a análise da métrica “Patentes publicadas” ( $X_{20}$  e  $X_{23}$ ), estrutura-se por meio de Tabelas Fato e Dimensões, adotando o esquema Floco de Neve.

#### **4.2.3 Passo 3: Definição de Métricas e KPIs**

Após a integração e modelagem dos dados, definiram-se métricas e KPIs relevantes para cada canal de PD&S, conforme as necessidades do usuário. Para o MVP, utilizaram-se métricas básicas, demonstrando a aplicabilidade da solução, embora possam

ser remodeladas para atender demandas específicas de pesquisadores, gestores de ICTs e tomadores de decisão. Utilizou-se a linguagem *Data Analysis Expressions* (DAX) no desenvolvimento das métricas e KPIs, linguagem esta, padrão para programação no Power BI, que permite cálculos complexos, a partir de dados de diversas tabelas do modelo de dados.

Como exemplo, uma métrica possível de ser calculada com DAX é a *quantidade de orientações* de pesquisa ( $X_3$ ) por ICT. Esse cálculo exigiu a criação de uma medida específica que organiza o número de  $X_3$  por ICT, sendo possível aplicar filtros conforme a necessidade, como por exemplo, restringir a análise a uma ou um conjunto de ICTs da REFPECT, considerando apenas aquelas submetidas por pesquisadores ( $X_2$ ) com dedicação exclusiva

Outro exemplo de KPI, calculado com DAX, representa a razão entre  $X_3$  e  $X_2$ . Esse indicador monitora a variação dessa razão ao longo do tempo ( $T_1$ ) e auxilia os gestores na decisão sobre a alocação de recursos, direcionamento de investimentos em pesquisa e *benchmarking*<sup>17</sup> entre instituições.

A codificação das métricas e KPIs em DAX pode ser alterada, conforme os filtros selecionados pelo usuário final, permitindo que os *scripts* sejam personalizados para fornecer informações específicas (Apêndice B, Figura B. 4 e Figura B. 12).

No presente estudo, a seleção de métricas, KPIs e gráficos foi criteriosa, contando com a experiência de dois ex-pró-reitores de Pesquisa, Desenvolvimento e Inovação, como consultores, adotando-se como referência os requisitos dos usuários e a análise de soluções públicas e privadas similares (Apêndice B, Quadro B. 17). Os indicadores desenvolvidos consideraram as demandas específicas de cada requisito e a disponibilidade dos dados coletados, mas alguns podem ser universalmente disponibilizados para instituições como MEC, SETEC, entre outras, permitindo a adoção desses indicadores em um contexto mais amplo, bem como o acompanhamento pelos órgãos de controle como TCU e CGU, em tempo real, facilitando aos gestores das ICTs a entrega de relatório, por exemplo.

Esta subseção ressalta exemplo dos possíveis resultados e suas importâncias, relevância e utilidade para os usuários, em conformidade com orientações normativas. A escolha apropriada de métricas e indicadores possibilita que os usuários adotem decisões

---

<sup>17</sup> *Benchmark* pode ser entendido como um processo de comparação entre as práticas, produtos e serviços de uma instituição com outras líderes em seu setor, visando identificar possíveis oportunidades de melhoria (Quevedo *et al.*, 2023).

orientadas e baseadas por/em evidências, contribuindo para o fomento e gestão das ICTs, no desenvolvimento de inovações e no suporte ao crescimento econômico.

#### 4.2.4 Passo 4: Criação do Painel Visual

Com os dados integrados e modelados, bem como as métricas e KPIs definidas, desenvolveu-se o *dashboard* no *Power BI*, estruturado em três módulos: **Amplio**, **Detalhado** e **Comparativo**, cada qual com filtros, métricas e KPIs desenhados para atender as necessidades dos usuários, conforme detalhado na Subseção 3.2.

Para otimizar a navegação e a interpretação das informações, planejou-se um *UX design*, simples, porém funcional e que demonstrasse no MVP o maior número possível de possibilidades e abrangência nos 3 módulos propostos.

O projeto gráfico deste MVP fundamentou-se em três pressupostos: i) a consideração do ecossistema da REFPCT, que possui uma identidade visual pré-estabelecida; ii) o público-alvo composto por sociedade, pesquisadores, gestores de ICTs e decisores; e iii) a flexibilidade do conceito de *UX design* para integrações futuras com outras soluções, conforme detalhado nas Subseções 3.2 e 3.6.

Para o projeto gráfico visual, adotou-se a identidade da REFPCT, aplicando as cores verde e vermelho conforme seu manual de identidade visual. Elas, conforme Pedrosa (1995) e Farina *et al.* (2011), foram escolhidas baseadas na psicodinâmica das cores, teoria que estuda seus efeitos na mente e nas emoções humanas. O verde, associado à natureza e ao equilíbrio, e o vermelho, vinculado à energia e à paixão, quando combinados, criam um contraste visual significativo, o que pode favorecer a hierarquia e o dinamismo no *design*.

A escolha dessas cores visa, portanto, a estabelecer um forte contraste visual que auxilie na interpretação das informações, bem como transmitir distintas emoções ao usuário e atender ao manual de identidade da Rede. No entanto, ressalta-se a importância de considerar as diferentes conotações culturais e sociais dessas cores em distintos contextos globais, permitindo assim que o *layout* do MVP seja adaptável a diversas realidades. O MVP está disponível em múltiplos idiomas, incluindo: Português e Inglês, com a possibilidade de inclusão de outros idiomas (Apêndice B, Figura B. 14).



### 4.3 Resultados Gerados pelo MVP Proposto

Nesta seção, são apresentados os principais resultados relacionados à organização de dados por meio da implementação do *UX design*, projetado para a plataforma proposta neste estudo, no formato de um MVP, utilizando as soluções apresentadas na Seção 3.

A seguir, são apresentados alguns exemplos dos resultados obtidos: i) a plataforma está disponível gratuitamente para acesso<sup>18</sup>; ii) atende a algumas demandas apresentadas por órgãos brasileiros de controle e planejamento de políticas vinculadas ao ecossistema de CT&I e PD&I (Brasil, 2019a; 2021b; 2022a, p. 50-53; 108-111; Marques *et al.*, 2022; Peixoto *et al.*, 2021); iii) nela é possível explorar seus filtros e soluções implementadas; iv) sua flexibilidade para integrar-se a outras soluções; e iv) a possibilidade de atualização dinâmica dos dados.

No campo científico, a plataforma busca resolver algumas das demandas frequentemente relatadas em estudos, sejam eles relacionados à análise acadêmica (Paz *et al.*, 2019), à visualização de dados no ensino superior (Azevedo *et al.*, 2021; Castillo *et al.*, 2018; Damyanov *et al.*, 2019; Di Tria *et al.*, 2015; Medina *et al.*, 2018; Paz *et al.*, 2019; Stewart *et al.*, 2022) e/ou à análise de indicadores de P&D (Azevedo *et al.*, 2021; Basso *et al.*, 2021; Chaves *et al.*, 2017; Fernandez *et al.*, 2017; Johnson *et al.*, 2022; Khatibi *et al.*, 2017; Menezes *et al.*, 2019; Miano *et al.*, 2020; Perucchi *et al.*, 2015; Ribeiro *et al.*, 2020; Santos *et al.*, 2015; Sierra *et al.*, 2018; Silva, 2015).

No campo de desenvolvimento tecnológico, a ferramenta, oriunda desta tese, resultou em um produto de propriedade intelectual tecnológico, consolidado através do registro de *software* junto ao INPI, sob o registro nº BR5120230010197.

No campo da prestação de serviços, os especialistas (Apêndice B, Quadro B. 3), apesar de proporem sugestões de melhoria, alcançam consenso quanto à capacidade inovadora do MVP em atender as demandas das ICTs e de Ministérios, como MEC e MCTI, por exemplo. Frise-se que as articulações com dois ministérios, e duas ICTs já foram inicializados.

No campo da implementação de soluções em BI, a plataforma, em tela, emprega técnicas de DW e *dashboard*, fundamentando-se nas teorias de Análise de Negócio, Visualização de Dados e da Tripla Hélice de Inovação. Essa abordagem demonstra seu potencial para acompanhar a performance das ICTs em relação aos canais de PD&S. Por

---

<sup>18</sup> Disponível em <https://prequest.websiteseuro.com/bi>, acessado em: 14/08/20203.

fim, conforme o perfil dos usuários, este MVP viabiliza a gestão de ações corretivas, de fomento, de auxílio, de pesquisa, de monitoramento e de consulta, bem como a solicitação e o atendimento de demandas dos SPP, GOV e sociedade, adaptando-se às necessidades e ao contexto de cada usuário.

Com relação à solução tecnológica, em tela, seu desenvolvimento resultou na criação de três módulos (**Ampla**, **Detalhado** e **Comparativo**). A disposição em 3 módulos permitiu uma organização analítica dos dados disponíveis. Baseou-se no nível de detalhe ou profundidade que os dados apresentam “granularidade” para gerar informações. A organização por granularidade permite ao usuário ajustar o nível de detalhamento dos dados, de acordo com suas necessidades específicas de análise e tomada de decisão. Portanto, cada módulo possibilita ao usuário obter os seguintes resultados:

- a) Personalizar análises: escolher o nível de detalhe que é mais relevante para sua questão ou problema, permitindo análises mais focadas e eficientes.
- b) Melhorar a visibilidade: alternar entre uma visão geral (baixa granularidade) e uma visão detalhada (alta granularidade) dos dados, facilitando a compreensão de tendências gerais ou de aspectos específicos.
- c) Otimizar o desempenho: evitar sobrecarga de informações, ao trabalhar com grandes volumes de dados, selecionando o nível de granularidade que é gerenciável e relevante para a tarefa em questão.
- d) Facilitar a comunicação: adaptar a apresentação dos dados para diferentes públicos, apresentando informações mais agregadas para um público não técnico e mais detalhadas para especialistas.
- e) Identificar padrões e anomalias: com a alta granularidade, é possível especificar padrões ou anomalias específicas que podem passar despercebidas em um nível mais agregado.
- f) Tomar decisões mais assertivas: utilizar o nível de detalhe apropriado para fundamentar decisões, seja para estratégias de longo prazo (usando dados de baixa granularidade) ou para ações táticas e operacionais imediatas (usando dados de alta granularidade).

A seguir, serão detalhados alguns dos resultados possíveis que cada um dos três módulos pode entregar ao usuário.

O módulo **Ampla** fornece uma visão geral das métricas selecionadas, utilizando gráficos de barra, nuvens de palavras, mapas de densidade e tabelas para representar os

resultados. Os filtros disponíveis permitem que os usuários ajustem as análises de acordo com suas necessidades, permitindo comparações entre diferentes ICTs ou com base em critérios específicos.

O módulo **Detalhado** segue uma abordagem semelhante, porém oferece análises mais detalhadas e granulares de cada métrica, aumentando a precisão da análise e auxiliando na identificação de ações corretivas. Essa granularidade permite que os usuários realizem análises aprofundadas de tendências, identifiquem padrões emergentes, comparem diferentes segmentos e períodos e examinem a distribuição dos dados em múltiplas dimensões. Além disso, a granularidade dos dados possibilita a identificação de gargalos e oportunidades, a avaliação da eficiência das políticas e estratégias em vigor e a elaboração de planos de ação específicos e personalizados para otimizar a performance do ecossistema RFEPCT por exemplo.

O último módulo, **Comparativo**, permite uma análise detalhada de, no mínimo, duas métricas simultaneamente, utilizando o método de *Z-score*<sup>19</sup> (Fávero *et al.*, 2020; Field, 2018; Hair *et al.*, 2019), para padronizar os valores e facilitar a comparação. A adoção da escala *Z-score* possibilita que os usuários possam identificar correlações e padrões ocultos entre as métricas, avaliar a influência de uma métrica em outra e analisar a interdependência entre indicadores.

Essa abordagem também possibilita a identificação de áreas de melhoria e a avaliação do desempenho relativo entre diferentes segmentos, instituições ou setores. Além disso, a comparação das métricas em escala *Z-score* permite o monitoramento das tendências ao longo do tempo e a análise do impacto das políticas e estratégias implementadas, por exemplo.

Os filtros disponíveis são semelhantes ao módulo **Amplio** e o gráfico de linhas se adapta dinamicamente aos filtros aplicados, proporcionando uma análise visual intuitiva, eficiente e visualmente dos resultados considerados significantes e com valores<sup>20</sup> de

---

<sup>19</sup> O *z-score*  $Z_i$  pode ser calculado subtraindo-se o valor individual da métrica (variável  $X_i$ ) da média  $\mu$  dessa mesma variável, dividindo sua diferença pelo desvio padrão  $\sigma$  dessa mesma variável:  $Z_i = (X_i - \mu) / \sigma$ .

<sup>20</sup> A expressão “ $-1 < r < -0,5$  ou  $0,5 < r < 1$ ” refere-se a um intervalo de valores possíveis para o coeficiente de correlação “ $r$ ” Pearson. Ele é uma medida que indica a força e a direção da relação linear entre duas variáveis. Varia de  $-1$  a  $1$ , onde  $-1$  indica uma correlação negativa perfeita. Um  $1$  indica uma correlação positiva perfeita, e o zero ( $0$ ), que não há relação linear entre as variáveis. A expressão “ $-1 < r < -0,5$  ou  $0,5 < r < 1$ ” indica a adoção de duas faixas de valores de correlação ( $r$ ) distintas: *a*) a faixa de valores entre  $-1$  e  $-0,5$  que representa uma correlação negativa; e *b*) A faixa de valores entre  $0,5$  e  $1$  representa uma correlação positiva. No MVP, consideraram-se apenas as correlações com uma intensidade moderada a forte, seja negativa ou positiva e significativas. Vale lembrar que correlação não implica necessariamente causalidade e que outros fatores podem estar envolvidos.

correlação de Pearson, entre  $-1 < r < -0,5$  ou  $0,5 < r < 1$  (Fávero *et al.*, 2020; Field, 2018; Hair *et al.*, 2019).

As métricas utilizadas nos três módulos representam a somatória ( $\Sigma$ ), considerando os filtros aplicados. Em todos os módulos, os atributos métricos considerados como dados institucionais são  $X_1$  e  $X_2$ ; os relacionados à pesquisa são  $X_3$ : $X_{11}$ ; os pertinentes à prestação de serviços são as variáveis  $X_{12}$ : $X_{17}$ ; e os vinculados ao desenvolvimento tecnológico são  $X_{18}$ : $X_{24}$ .

A Figura 4a mostra os ícones que concedem acesso aos módulos (Apêndice F, Figura F. 1: subitem “g”). A Figura 4b representa o ícone que, ao clicar, desfaz os filtros aplicados (Apêndice F, Figura F. 1: subitem “a”). A Figura 4c exibe o ícone que, ao clicar, retorna à página de apresentação geral da plataforma (Apêndice F, Figura F. 1: subitem “a”). A Figura 4d apresenta a lista de variáveis disponíveis para consulta nos módulos **Amplio** e **Detalhado** (Apêndice B, Figura B. 15 e Figura B. 16). Por fim, A Figura 4e mostra a lista de variáveis disponíveis para consulta no módulo **Comparativo** (Apêndice B, Figura B. 17 e Apêndice F, Figura F. 1: subitem “a”).

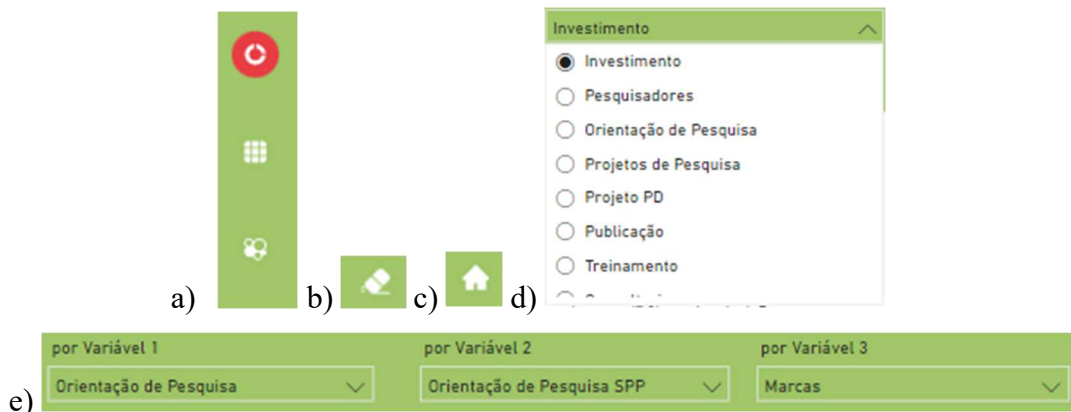


Figura 4: Visão geral das funcionalidades do dashboard MVP em diferentes módulos.

A Figura 4 ilustra as funcionalidades gerais da plataforma proposta em todos os três módulos, destacando sua versatilidade. Os ícones que possuem cores indicativas de seu *status* (vermelho para ativo, cinza para desativado e branco para habilitado), permitem navegar entre os três módulos (a), desfazer filtros aplicados (b) e retornar à página de apresentação geral (c). A figura também mostra a lista suspensa para selecionar as métricas/variáveis nos módulos **Amplio** e **Detalhado** (d), possibilitando análises univariadas a partir dos filtros disponíveis e selecionados, conforme os critérios do usuário. Além disso, o painel exibe a lista suspensa para consultar até três variáveis

simultaneamente no módulo **Comparativo** (e), viabilizando análises multivariadas com base nos filtros selecionados.

As variáveis disponíveis nesta versão MVP podem ser acessadas com um simples “clique” sobre a opção desejada (Figura 4: subitem "a", "b", "c", "d" e "e"), proporcionando um *UX design* adequado às condições, ao tempo e aos recursos disponíveis.

O *UX design* proposto (Apêndice B, Figura B. 15-Figura B. 17 e Apêndice F, Figura F. 1) e o modelo de BI adotado (Figura 2) foram desenvolvidos visando flexibilidade e adaptabilidade a novos contextos, conforme constatado durante levantamento de requisitos e *feedbacks* dos especialistas. Dessa forma, foi possível evitar a complexidade excessiva, que poderia tornar a manutenção, adaptabilidade e integração custosas, conforme destacado por Erdos *et al.* (2019).

O MVP foi cuidadosamente planejado para garantir uma experiência de usuário intuitiva, organizada e eficiente, permitindo fácil acesso às informações e às ferramentas necessárias. Para uma análise detalhada da simulação realizada com o *usuário W* de uma ICT, que supostamente adquiriu a ferramenta, em tela, consulte Apêndice H.

Assim, ao considerar a grande quantidade de cruzamentos de dados direcionados ao monitoramento, performance, publicitação e transparência dos resultados, o *usuário W* identificou oportunidades para criar ações, diretrizes e políticas institucionais que potencializem as interações da sua ICT com o SPP e, adicionalmente, o GOV. A partir dessas informações, o *usuário W* está agora em posição de implementar as seguintes ações específicas, diretrizes e política institucional para a sua ICT; seguem alguns exemplos:

#### **Ações Específicas:**

- Estabelecer parcerias estratégicas com o SPP e GOV para aumentar o número de *projetos* de *P&D* colaborativos, incentivando a troca de conhecimento e a sinergia entre pesquisadores e profissionais;
- Oferecer programas de treinamento e capacitação para pesquisadores e profissionais envolvidos em *projetos* de *P&D*, a fim de melhorar suas habilidades e conhecimentos em metodologias e técnicas avançadas de pesquisa; e
- Estimular a participação em eventos acadêmicos e profissionais, como conferências e *workshops*, para promover a disseminação de resultados de

pesquisa e fomentar a colaboração entre pesquisadores e profissionais de diferentes setores.

**Diretrizes:**

- Desenvolver um plano de ação detalhado que estabeleça objetivos claros, metas e prazos para a implementação das ações específicas, garantindo que todos os envolvidos estejam alinhados e comprometidos com a realização dessas ações; e
- Estabelecer mecanismos de acompanhamento e avaliação das ações implementadas, permitindo identificar o progresso em direção aos objetivos e metas estabelecidos, bem como ajustar as ações conforme necessário para garantir a eficácia e o impacto desejado.

**Política Institucional:**

- Criar uma política institucional que promova a colaboração e a cooperação entre ICTs- SPP - GOV, facilitando o compartilhamento de recursos, conhecimento e experiências, e incentivando a inovação conjunta e a integração das atividades de pesquisa e desenvolvimento entre os diferentes ecossistemas. Isso pode incluir a criação de comitês de coordenação conjunta, a realização de eventos e reuniões regulares, e a implementação de processos formais de comunicação e cooperação.

Ao analisar os resultados obtidos no estudo de caso hipotético realizado pelo “usuário W” (Apêndice H) e as avaliações reais dos especialistas (Apêndice B, Quadro B. 3), foi possível constatar a validade, relevância e potencialidade, deste MVP, para auxiliar os gestores na tomada de decisões e os usuários na análise em ecossistemas de ICTs.

Nove especialistas, representando os três setores da TH, foram convidados para avaliar a proposta tecnológica detalhada neste estudo, cujo foco foi validar o objetivo geral, a questão de pesquisa e a hipótese apresentada. Eles destacaram o potencial do MVP, apesar de apontarem a necessidade de ajustes na interface e na definição de indicadores e reconhecem a importância de o expandir para outras áreas e instituições, demonstrando um consenso na percepção de que ele tem muito a oferecer. São descritos

abaixo os resultados obtidos das análises dos especialistas para a versão MVP proposta neste trabalho:

### **Tópico 1 - Relevância e Potencialidade:**

- A ferramenta traz indicadores importantes sobre a performance dos institutos federais de ensino (especialista 1);
- A ferramenta concentra em uma única plataforma dados coletados de várias fontes (especialista 2);
- A ferramenta apresenta utilidade para a Rede Federal de Ensino Profissional na área de pesquisa e inovação (especialista 3);
- A ferramenta apresenta relevância para gestores de políticas públicas e gestores institucionais (especialista 5);
- O MVP é útil para fomentar políticas públicas e potenciais melhorias no sistema de ensino (especialista 6);
- A ferramenta atende ao objeto proposto, sendo considerada uma importante ferramenta de auxílio à gestão de políticas de PD&I e com grande potencial de uso (especialista 7);
- A ferramenta é concebida como de grande relevância, permitindo aos gestores (reitores e pró-reitores) ter uma visão macro de suas instituições na vertente de PD&I (especialista 8);
- O produto é relevante para a gestão e acompanhamento de ações/processos de pesquisa, atendendo os objetivos a que se propõe (especialista 9).

### **Tópico 2 - Potencial de Mercado:**

- A ferramenta pode ser customizada para aferir e monitorar ações de outras instituições (especialista 1);
- A ferramenta pode ser customizada para contemplar outras instituições do ecossistema de CT&I do Brasil (especialista 2);
- A ferramenta apresenta viabilidade comercial devido à relevância do problema e qualidade da solução (especialista 5);
- O produto possui viabilidade de comercialização com baixo custo (especialista 6);

- A ferramenta é percebida como pouco viável comercialmente. Contudo, na perspectiva de uma venda coletiva. Como a aquisição pela SETEC/MEC para uso da rede, apresenta boa viabilidade comercial (especialista 7);
- A ferramenta é viável para uma transferência de tecnologia, através de licenciamento, uma vez que cada instituição que deseje utilizá-la precisará ter um contrato para adequação das funcionalidades à sua realidade (especialista 8);
- O produto apresenta potencial para viabilidade comercial (especialista 9).

### **Tópico 3 - Usabilidade e Interface:**

- A ferramenta apresenta design amigável e boa navegabilidade (especialista 2);
- Apresenta necessidade de melhorias na interface e definição de indicadores (especialista 5);
- Apresenta necessidade de um manual de utilização e mais abas laterais de navegação (especialista 6);
- Sugere-se incluir nela opções de dados e densidade mais qualitativos na ferramenta, como número de projetos por número de docentes e valor de investimento por número de projetos (especialista 7);
- Nela, há falta de um ambiente que permita visualizar especificamente uma instituição de interesse e gráficos comparativos entre instituições. É sugerida a inclusão de funcionalidades para responder a perguntas comparativas entre instituições (especialista 8);
- Nela, não há observações específicas sobre usabilidade e interface no parecer do especialista (especialista 9).

### **Tópico 4 - Aplicabilidade para Gestores e Tomada de Decisão:**

- A ferramenta instrui os gestores sobre melhores políticas a serem implementadas (especialista 1);
- Auxilia na tomada de decisão, formulação e avaliação de políticas (especialista 2);



- Possui relevância do sistema para tomada de decisão e ampliação das áreas de pesquisa (especialista 3);
- Facilita a análise de dados para gestores do setor produtivo e geração de projetos de empreendedorismo (especialista 4);
- É avaliada como um recurso significativo para auxiliar gestores na tomada de decisões, potencializando o uso de recursos através de uma avaliação precisa e embasada (especialista 7);
- Tem grande potencial para ser adotada por gestores, podendo ser utilizada no contexto de ICT com os devidos ajustes. É ressaltada a necessidade de ajustes para permitir gráficos comparativos e correlações entre variáveis iniciais e resultados (especialista 8);
- Do ponto de vista da gestão promove um excelente mapeamento de informações importantes que pode facilitar a tomada de decisões e direcionamento de ações diretas de aprimoramento dos processos de pesquisa, desenvolvimento e inovação das instituições (especialista 9).

#### **Tópico 5 - Sugestões e Melhorias:**

- A ferramenta ajusta-se às necessidades de agências de fomento à pesquisa (especialista 1);
- Necessita melhorar suas características intuitivas dos botões e inserir manual *online* (especialista 2);
- Necessita incluir informações de referência das fontes e bancos de dados, modelo de maturidade de inovação e caracterizar fonte de recurso (especialista 3);
- Necessita a integração de API de serviços de entrega de dados e painel de gerenciamento de usuários (especialista 4);
- Necessita melhorar a organização da interface e inclusão de maior variedade de ICTs (especialista 5);
- Ajustar o tempo de resposta da ferramenta, pois é considerado alto para o tipo de aplicação, sendo sugerida uma otimização do sistema para respostas mais rápidas. Questiona-se a possibilidade de um módulo de importação de dados da plataforma Lattes (especialista 7);

- Revisar e melhorar o entendimento das categorias de pesquisa, como a diferenciação entre "orientação de pesquisa" e "estágio supervisionado". É sugerida a revisão de parâmetros, como o tempo médio para publicação de patentes e a conexão entre seleção de variável e os resultados exibidos (especialista 8).

#### **Tópico 6 - Comparação com Outras Soluções:**

- Efetividade em comparação com o sistema SparkX do Instituto Stela (especialista 1);
- Comparação com iniciativa semelhante da SECTI-BA e FAPESB (especialista 2);
- Vantagem de acesso público e gratuito em comparação com Stela Experta (especialista 4);
- Excelente inovação em relação a outras soluções (especialista 6);
- O produto tecnológico encontra-se bem amplo, apresentando informações que outras soluções no mesmo sentido não apresentam. O especialista destaca, como exemplo, o investimento destinado a cada instituição, que é uma informação fundamental para a gestão e tomada de decisão relacionadas a ações/processos que envolvem a Pesquisa e a Inovação (especialista 8);

Considerar os resultados teóricos, o estudo de caso hipotético e as avaliações dos nove (9) especialistas permite concluir que o MVP tem potencial para auxiliar os usuários na tomada de decisões baseada em dados, atingindo, assim, o objetivo geral deste trabalho.

Implementar as melhorias sugeridas pelos especialistas permitirá oferecer, em uma nova versão, mais opções para análise dos dados, garantindo maior alcance e impacto no futuro. O MVP, ao promover transparência, acessibilidade e facilitação de auditorias dos dados públicos, pretende fortalecer o ecossistema de pesquisa e desenvolvimento nacional, contribuindo decisivamente para o avanço da CT&I e PD&I no Brasil.

#### 4.4 Resultados Obtidos Sobre a Questão de Pesquisa e Hipótese Investigada

No estudo de caso realizado no período entre 1 de janeiro de 2008 e 31 de dezembro de 2018, procedeu-se à avaliação da eficácia de uma solução tecnológica proposta, visando não apenas validar sua relevância e efetividade, mas também identificar possíveis ajustes ou melhorias.

Os resultados permitiram tanto validar os tópicos, supracitados, como também aprimorar o MVP à capacidade de resposta dos gestores das ICTs às demandas institucionais, produtivas e governamentais, incluindo aquelas provenientes dos órgãos de controle governamentais e indiretamente reduzir o hiato entre a teoria consolidada e a pesquisa empírica atual, sem ignorar a complexidade intrínseca do SNI e os diversos ecossistemas das ICTs.

Mesmo integrando uma rede unificada como a RFEPCT, é essencial considerar as individualidades e heterogeneidades presentes. Isso se reflete não apenas na constituição destas entidades, mas também na maneira como interagem com os canais de PD&S, influenciando diretamente os resultados alcançados. A seguir, são apresentados os resultados das avaliações do MVP aqui proposto, com vistas aos atributos e às potencialidades da solução pelos especialistas (Subseção 2.3 e Apêndice B, Quadro B. 3).

Os especialistas foram unânimes em destacar a relevância e efetividade da solução tecnológica proposta. Entre os *feedbacks* recebidos, ressaltou-se a importância da ferramenta para a tomada de decisões estratégicas e a gestão otimizada dos referidos canais. Sugeriu-se, predominantemente, a realização de ajustes específicos e a otimização de determinadas características, a fim de alinhar a ferramenta às necessidades das FAPs e facilitar sua integração com outras plataformas já existentes (Especialista 1).

Ao analisar individualmente as considerações dos especialistas, observou-se que a customização da ferramenta para outros ecossistemas foi um ponto abordado, destacando-se a potencialidade de expansão para futuros clientes (Especialista 2). Além disso, salientou-se a necessidade de melhorias em características específicas, incluindo a definição de indicadores mais precisos e a implementação de gráficos comparativos (Especialista 3, 5 e 8).

No que concerne à usabilidade da ferramenta, foram propostos, como estratégias para tornar a solução mais acessível e funcional, a implementação de um manual de utilização e o ajuste na interface (Especialista 5 e 6).

Foi ressaltado, ainda, o potencial da ferramenta para a gestão otimizada de políticas, indicando-se a necessidade de otimizações que permitam uma inserção mais harmoniosa de novas variáveis e parâmetros (Especialista 7 e 9).

Em suma, a análise dos *feedbacks* realizada permitiu, validar a eficácia da solução tecnológica proposta e identificar áreas prioritárias para ajustes e melhorias.

O consenso entre os especialistas ressalta a relevância e a utilidade pública do MVP, projetando um cenário positivo para sua implementação, desde que acompanhada das devidas otimizações sugeridas.

#### 4.5 Resultados Descritivos da RFEPCT e da Revisão da Literatura

Os resultados e a capacidade de inovação institucional da RFEPCT, após 10 anos de funcionamento, representam um componente fundamental, refletindo-se nos resultados entre as ICTs. Esses aspectos são essenciais para compreender a contribuição dessas instituições ao avanço da CT&I e da PD&I no país. Eles e suas respectivas análises referem-se a um compilado de dados colhidos dos canais de PD&S das ICTs entre 2008 e 2018.

Os resultados, oriundos da análise descritiva dos dados da RFEPCT (Apêndice B, Quadro B. 13 e Figura B. 3, Tabela B. 3 e Tabela B. 4), estão disponíveis no MVP<sup>21</sup>, através dos módulos **Amplio**, **Detalhado** e **Comparativo**. O conjunto de dados PD&S adotado captura registros de diversas ICTs ao longo de uma década, fornecendo detalhamento de 24 métricas variadas ( $X_1$ : $X_{24}$ ), abrangendo desde investimentos ( $X_1$ ) a programas de computador em interação com os SPP ( $X_{24}$ ).

As métricas englobam dados relacionados ao âmbito Institucional (GOV), de Pesquisa (ICTs e ICTs-SPP), de Serviços (ICTs e ICTs-SPP) e de Desenvolvimento Tecnológico (ICTs e ICTs-SPP), variando entre: quatro (4) categóricas, uma (1) contínua e vinte e três (23) discretas (Apêndice B, Quadro B. 14).

Investimento Público em P&D ( $X_1$ ): essa variável pode ser analisada em relação à quantidade de recursos financeiros captados e aplicados para o fomento dos canais de PD&S, permitindo avaliar o impacto direto desses investimentos no avanço da pesquisa, desenvolvimento e inovação, a eficácia das políticas públicas implementadas e a

---

<sup>21</sup> Disponível em <https://prequest.websiteseuro.com/bi>, acessado em: 14/08/20203.

compreensão de que forma tais investimentos podem impulsionar avanços tecnológicos significativos, contribuindo para o crescimento socioeconômico regional e sustentável.

Pesquisadores ( $X_2$ ): essa variável pode ser analisada em relação à titularidade e anos de serviço, permitindo avaliar a experiência e a qualificação dos pesquisadores envolvidos nas atividades de pesquisa e desenvolvimento.

Orientações ( $X_3$ ) e Projetos de pesquisa ( $X_4$ ): essa variável pode ajudar a quantificar a capacidade das instituições de conduzir pesquisas de alta qualidade e inovadoras, bem como a interação com os setores produtivos.

Projetos de Pesquisa SPP ( $X_8$ ) e Cooperação ( $X_9$ ): essa variável pode ser usada para medir a colaboração entre as instituições e os setores produtivos, podendo indicar o nível de inovação alcançado.

Treinamento ( $X_{10}$ ) e Consultoria ( $X_{11}$ ): essas variáveis podem ser usadas para avaliar a capacidade das instituições de transferir conhecimento e tecnologia para os setores produtivos, bem como a qualidade do suporte fornecido aos pesquisadores e estudantes.

Patentes ( $X_{20}$ ) e Programa de Computador ( $X_{21}$ ): essas variáveis podem ser usadas para avaliar a capacidade de inovação das instituições e a aplicação prática dos resultados de pesquisa.

As Tabelas 4 e 5 relatam o somatório dos dados originais ( $X_i$ ) entre 2008 e 2018 de todos os vinte e oito parâmetros observados para cada ICT e suas estatísticas descritivas, respectivamente. Os parâmetros foram: ( $W_1$ ) Região, ( $W_2$ ) UF, ( $W_3$ ) IPE, ( $X_1$ ) Investimento Público em P&D, ( $X_2$ ) Pesquisadores, ( $X_3$ ) Orientações de Pesquisa, ( $X_4$ ) Projetos de pesquisa, ( $X_5$ ) Projetos de P&D desenvolvidos exclusivamente por uma ou mais IPE, ( $X_6$ ) Publicações Científicas, ( $X_7$ ) Orientação de pesquisa financiada pelo SPP, ( $X_8$ ) Projetos de pesquisas em cooperação, ( $X_9$ ) Projetos de pesquisas financiados, ( $X_{10}$ ) Treinamentos para IPE-GOV, ( $X_{11}$ ) Consultoria para IPE-GOV, ( $X_{12}$ ) Projetos de extensão para IPE-GOV, ( $X_{13}$ ) Uso de laboratórios e equipamentos etc. - Outras Atividades. para IPE-GOV, ( $X_{14}$ ) Treinamentos para SPP, ( $X_{15}$ ) Consultoria para SPP, ( $X_{16}$ ) Projetos de extensão para SPP, ( $X_{17}$ ) Uso de laboratórios e equipamentos etc. - Outras Atividades. para SPP, ( $X_{18}$ ) Marcas, ( $X_{19}$ ) Desenhos Industriais, ( $X_{20}$ ) Patentes Publicadas, ( $X_{21}$ ) Programas de computador, ( $X_{22}$ ) Desenhos industriais IPE-SPP, ( $X_{23}$ ) Patentes Publicadas IPE-SPP, ( $X_{24}$ ) Programas de computador IPE-SPP.

A seguir, serão relatadas as principais descobertas obtidas neste estudo. Entre elas destacam-se: Investimentos e Distribuições, Barreiras e Facilitadores, Correlações,

Indicadores de Desempenho e, por fim, a Análise Descritiva de alguns dados considerados mais relevantes.

#### **4.5.1 Sinergia de Dados: Resultados da Análise Descritiva e Revisão da Literatura**

Conforme detalhado nas Seções 2 e 3, para potencializar a capacidade de resposta às necessidades institucionais, produtivas e governamentais - inclusive aquelas demandadas pelos órgãos de controle - é crucial analisar e acompanhar continuamente os resultados, e é nesse contexto que o MVP aqui proposto pode auxiliar o SNI a integrar e otimizar as estratégias e operações das ICTs, bem como a fomentar a inovação e o desenvolvimento sustentável, através da articulação eficaz de iniciativas de CT&I e PD&I.

O MVP aqui proposto é aplicável, viável, relevante e singular (subseção 4.4), apresentando uma abordagem que vai além de um modelo conceitual (subseção 4.2), permitindo análises longitudinais, transversais e integração com outras plataformas (subseção 3.5.1).

A aplicação da técnica de BI é relevante no campo teórico para entender as complexidades e interações entre os atores envolvidos e para aprimorar a tomada de decisões e políticas em PD&S, promovendo inovação e desenvolvimento (Albats *et al.*, 2018; Etzkowitz *et al.*, 1995; 2000; Rabova *et al.*, 2014).

Sistemas baseados em BI, aplicados no campo das ICTs, podem melhorar a eficiência e eficácia de sua gestão organizacional, contribuindo para potencializar a qualidade e relevância da entrega dos seus resultados por meio de seus canais, em especial, os de PD&S (Azevedo *et al.*, 2021; Di Tria *et al.*, 2015).

Constatou-se a necessidade de:

- considerar os fatores organizacionais, culturais e infra estruturais e nas implementações tecnológicas para o desenvolvimento de sistemas de BI (Figueiredo *et al.*, 2020; Santos, Thiago Rizzi *et al.*, 2021; Skyrius, 2021).
- conduzir análises robustas e abrangentes para ajustes contínuos de estratégias, integrando múltiplas fontes de dados e padronizando KPIs (Azevedo *et al.*, 2021; Figueiredo *et al.*, 2020; Wang *et al.*, 2013).
- desenvolver mais estudos interdisciplinares e integrados que considerem as especificidades e necessidades das ICTs para atender às demandas dos

setores envolvidos no processo de inovação (Etzkowitz *et al.*, 1995; 2000; Leydesdorff *et al.*, 1996; 1998; Martins *et al.*, 2022; Rantala *et al.*, 2018).

- desenvolver soluções tecnológicas que potencializem as ICTs, pois são elas, na maioria dos países, responsáveis pela produção científica nacional. No Brasil, em especial, elas são responsáveis por mais de 95% dessa produção (Analytics, 2018; 2019).
- consultar as bases do INPI e do Ministério da Agricultura e Pecuária, por outros tipos de registros industriais, entre 2008 e 2018, como indicações geográficas, transferências tecnológicas, topografias de circuitos integrados ou registro de cultivares; após consultas, tais registros não foram localizados até o período analisado.

#### 4.5.2 Investimentos e Distribuições

O estudo minucioso realizado na Tabela B. 3 (Apêndice B) permitiu identificar padrões e discrepâncias significativas entre as ICTs da RFEPCT entre 2008 e 2018.

O IFPB na Paraíba, por exemplo, destacou-se como principal destinatário dos investimentos em P&D na região Nordeste, seguido pelas regiões Sudeste e Sul.

As instituições CEFET/MG e IFPB destacaram-se pelos altos valores em determinados indicadores, representando pontos fora da curva, no intervalo temporal entre 2008 e 2018, analisado.

Na região Sudeste, constatou-se uma concentração considerável de recursos no CEFET/MG, totalizando mais de 54 milhões, enquanto o IFSP apresentou a maior quantidade de projetos de pesquisa ( $X_4$ ), alcançando o número de 2298 projetos.

No Nordeste, o IFPB sobressaiu-se tanto em termos de recursos, ultrapassando os 81 milhões, quanto no número de pesquisadores ( $X_2$ ), contabilizando 1181 pessoas. Na mesma região, o IFCE exibiu a maior quantidade de dados variáveis positivos, demonstrando uma pluralidade significativa em seus canais de PD&S.

Na região Sul, o IFC destacou-se pelo volume significativo de  $X_2$ , somando 1014 indivíduos, enquanto o IFFAR registrou mais de 3,5 milhões em  $X_1$ .

Na análise detalhada dos dados, constatou-se que algumas instituições apresentaram zeros (0) em muitas das variáveis analisadas, indicando a necessidade de uma ferramenta que auxilie os gestores também para entender as possíveis ausências de registros em suas ICTs, além das outras possíveis soluções que o MVP busca atender.

### 4.5.3 Correlações

A Tabela B. 5 (Apêndice B) apresenta as correlações entre diversas métricas, excluindo-se a *outlier* UTFPR, iterativamente utilizando o MVP *online*, pode-se incluir a UTFPR e realizar outros filtros. Para o exemplo, proposto, foi calculado o coeficiente de correlação entre cada par de métricas, indicado pelos valores numéricos na tabela. O coeficiente de correlação varia de -1 a +1, no qual +1 indica uma correlação positiva perfeita, -1 indica uma correlação negativa perfeita e zero (0) indica nenhuma correlação linear.

Para cada coeficiente de correlação, foi realizado um teste de significância para determinar se a correlação é estatisticamente significativa. A significância é indicada pelos *p*-valores entre parênteses, e marcas como " , " e " representam níveis de significância (geralmente, . para  $p < 0,1$ , \* para  $p < 0,05$ , \*\* para  $p < 0,01$ , e \*\*\* para  $p < 0,001$ ).

O nível de significância para os testes estatísticos foi, presumivelmente, de 0,05, o que permitiu a identificação de correlações significativas entre as métricas. Observa-se que correlação não implica causalidade; ela apenas indica que existe uma relação linear entre duas variáveis, detalhes consultar Fávero *et al.* (2020), Hair *et al.* (2019) e Field (2018).

Dentre as correlações significativas, destacam-se  $corr(X_3, X_2; r(41) = 0,82, p < 0,001)$ , onde " $r(41)$ " indica o coeficiente de correlação de Pearson ( $r$ ) com 41 graus de liberdade;  $corr(X_5, X_2; r(41) = 0,93; p < 0,001)$ ,  $corr(X_3, X_6; r(41) = 0,90; p < 0,001)$ ,  $corr(X_{12}, X_{15}; r(41) = 0,89; p < 0,001)$ ,  $corr(X_{19}, X_{22}; r(41) = 0,87; p < 0,001)$  e  $corr(X_{23}, X_{24}; r(41) = 0,89; p < 0,001)$ .

Dentre as correlações moderadas, destacam-se  $corr(X_3, X_1; r(41) = 0,42; p < 0,01)$ ,  $corr(X_9, X_3; r(41) = 0,67; p < 0,001)$  e  $corr(X_6, X_{20}; r(41) = 0,53; p < 0,001)$ .

Dentre as correlações fracas, destacam-se  $corr(X_2, X_1; r(41) = 0,25; p < 0,11)$ ,  $corr(X_4, X_1; r(41) = 0,24; p < 0,13)$  e  $corr(X_7, X_1; r(41) = 0,30; p < 0,06)$ , que também estão de acordo com os *p*-valores, acima de 0,05.



#### 4.5.4 Indicadores de Desempenho

Azevedo *et al.* (2021), Figueiredo *et al.* (2020) e Wang *et al.* (2013) reforçam a necessidade de análises robustas e abrangentes, incluindo estudos longitudinais, integração de múltiplas fontes de dados e padronização de KPIs.

Os KPIs iniciais sugeridos para o MVP buscaram padronizar diversas médias, durante a análise de requisitos, tais como o número de patentes publicadas em colaboração com os SPP ( $X_{23}$ ) registrando 0,21 de medida; de atividades de interação com SPP ( $X_{17}$ ), por pesquisador, registrando 0,13; valor de investimentos em P&D ( $X_1$ ), por pesquisador, totalizando R\$ 13.690,97 ou o número de Projetos de Pesquisa em interação SPP ( $X_8$ ), por pesquisador, registrando 0,001 como medida.

#### 4.5.5 Análise Descritiva

Para relatar o resultado relacionado aos dados contínuos, como o valor de  $X_1$  foi utilizada a média. Para os dados discretos, foi utilizada a mediana e a moda, por serem uma abordagem apropriada para descrever a tendência central de dados discretos, como o número de pesquisadores ( $X_2$ ). A mediana é uma medida robusta que não é afetada por valores extremos e a moda indica o valor que ocorre com maior frequência nos dados.

A análise revelou que, ao longo da década, o  $X_1$  médio nas ICTs foi de R\$ 134.823.349,26, com um desvio padrão de R\$ 18.420.204,76, indicando um coeficiente de variação de  $X_1$  entre as instituições de 136,62%, o que indica uma alta variabilidade em relação à média. O Gráfico B. 1 (Apêndice B) apresenta flutuações nos valores de  $X_1$ , variando de zero a aproximadamente R\$ 104,5 milhões, no decorrer dos anos de 2008 a 2018.

As ICTs registraram uma mediana de 911 para  $X_2$  e uma moda de 1.016 durante o período analisado. O IFSP teve o maior número de  $X_2$ , enquanto o IFRR apresentou o menor contingente.

As *orientações de pesquisa* ( $X_3$ ) e *projetos de pesquisa* ( $X_4$ ) apresentaram, no período analisado, uma mediana de 5.960 e 3.345, e moda de 50.110 e 12.108, respectivamente.

As ICTs que apresentam um número elevado de  $X_3$ , foram: UTFPR, IFCE, IFES, IFRJ, IFRN, CEFET/MG, IFSC, IFGO, IFBA, IFGOIANO, IFSP, IFSUL, IFMA, IFRS, IFC, IFMT, IFMG, IFPB, IFPR, IFPI e IFFAR. Já as ICTs com um número elevado de

X<sub>4</sub>, acima da mediana, foram: UTFPR, IFCE, IFES, IFRJ, IFRN, IFSC, IFGO, IFBA, IFTM, IFGOIANO, IFSP, IFSUL, IFMA, IFRS, IFC, IFMT, IFPA, IFPB, IFPR, IFPI e IFFAR.

As *publicações* (X<sub>6</sub>) por ICTs apresentaram uma mediana de 1.810, com uma moda de 16.445, no período correspondente. A UTFPR liderou em termos de X<sub>6</sub>, enquanto o IFRR teve o menor número.

As *patentes* publicadas (X<sub>20</sub>) tiveram uma mediana de 4 e uma moda de zero. CEFET/MG e IFFAR lideram como depositantes no INPI, enquanto IFAM e IFAL, entre outras ICTs, não haviam registrado patentes até o período considerado.

Por fim, alguns resultados obtidos no módulo '**Detalhado**' do MVP, referentes às *orientações de pesquisa* (X<sub>3</sub> e X<sub>7</sub>) dentro do âmbito da RFEPCT, considerando o resultado temporal obtido entre 2008:2018.

- X<sub>3</sub> são majoritariamente conduzidas por pesquisadores com dedicação exclusiva (92,55%);
- X<sub>3</sub> são majoritariamente orientadas por mestres (53,08%);
- X<sub>3</sub> são majoritariamente realizadas na área das engenharias (24,87%);
- X<sub>3</sub> são majoritariamente desenvolvidas para serem aplicadas ao setor CNAE P, relacionado à educação (38,24%) (Quadro B. 16, Apêndice B);
- 1,5% das X<sub>3</sub> são consideradas pelos pesquisadores como relevantes;
- 25,59% das X<sub>3</sub> foram financiadas por bolsa;
- X<sub>3</sub> são majoritariamente do tipo outras (31,43%);
- X<sub>7</sub> representam 1,67% de interação com SPP;
- 100% das X<sub>7</sub> foram financiadas por bolsa;
- 1,67% das X<sub>7</sub> são consideradas pelos pesquisadores como relevantes;
- X<sub>7</sub> são majoritariamente do tipo iniciação científica (56,48%);
- X<sub>7</sub> são majoritariamente realizadas na área das ciências exatas e da terra (23,46%);
- X<sub>7</sub> são majoritariamente desenvolvidas para serem aplicadas ao setor CNAE P, relacionado à educação (43,99%) (Quadro B. 16, Apêndice B);
- X<sub>7</sub> são majoritariamente financiadas pelas FAPs (22,25%);
- 0,68% das X<sub>7</sub> foram financiadas pela Petrobras e pela Agência Nacional de Petróleo.

- $X_7$  são majoritariamente desenvolvidas na região Sudeste, em especial, nas unidades da federação RJ e ES.

#### 4.5.6 Análise de Outliers (Uni)Multivariada

Durante o período analisado, identificou-se a UTFPR como um *outlier*, tanto em análises univariadas quanto multivariadas, constituindo-se num padrão que se manteve estável ao longo dos anos considerados. Ressalta-se que essa condição se verificou consistentemente, independentemente das diferentes abordagens empregadas.

A mesma observação não pôde ser aplicada a outras instituições, como o CPPI e o IFPA, que se destacaram como possíveis *outlier* em momentos isolados. No entanto, a análise consolidada dos dados, coletados entre 2008 e 2018, não sustentou essa classificação para o período completo.

### 4.6 Resultados da Mineração dos Dados da RFEPCT

A *Data Mining* (Apêndice C e Apêndice B, Tabela B. 7) envolve o uso de algoritmos para identificar padrões em grandes conjuntos de dados, enquanto a BI refere-se ao processo mais abrangente de coleta, armazenamento, análise e interpretação de dados de negócios para auxiliar na tomada de decisões.

Os métodos de agrupamento definidos como HC e NHC e, a EFA por PCA (Apêndice C, Subseção C.1 e C.2) são ferramentas que auxiliam na identificação de grupos dentro dos dados ou na redução de sua dimensionalidade, respectivamente.

Estas técnicas, ao serem aplicadas aos dados disponíveis, podem proporcionar *insights* valiosos para usuários que utilizam ferramentas baseadas em BI, como, por exemplo, os tomadores de decisão.

#### 4.6.1 Resultados das Análises HC e NHC

A análise HC (Subseção 3.5.9.1 e Apêndice C, Subseção C.1) visa compreender a taxionomia hierárquica ao agrupar observações conforme suas semelhanças. A análise NHC, utilizando técnicas de otimização, tem por finalidade identificar subconjuntos naturais em um conjunto amplo de dados. Nestas análises concentrou-se no intervalo temporal entre 2008 e 2018, a partir de duas abordagens distintas para entender as relações

entre um conjunto heterogêneo de ICTs que formam a RFEPCT, a saber Abordagem 1 e 2 (Apêndice B, Figura B. 18 e Figura B. 19).

Para a apresentação dos resultados da Abordagem 1 (Apêndice B, Quadro B. 19), durante a análise HC, empregaram-se todas as métricas ( $X_1:X_{24}$ ), sem levar em consideração a contribuição individual de cada uma para a formação dos agrupamentos; nesse contexto, optou-se pela exclusão das observações associadas à UTFPR e ao CPII, identificadas como *outliers*. Em contrapartida, a Abordagem 2 (Apêndice B, Quadro B. 20) focou exclusivamente nas métricas que influenciaram de maneira direta a construção dos grupos, desconsiderando unicamente a observação referente à UTFPR, reconhecida como outlier.

No que concerne à análise NHC, manteve-se o critério de seleção de métricas conforme estabelecido para cada abordagem (Apêndice B, Quadro B. 21 e Quadro B. 22) previamente descrita. Entretanto, é importante ressaltar que, independentemente da abordagem adotada, apenas a UTFPR foi categorizada como outlier, seja sob uma perspectiva univariada ou multivariada.

Constatou-se que após a aplicação da ANOVA *one way*, às análises HC e NHC, os resultados foram similares (Apêndice B, Quadro B. 18, Figura B. 18 e Figura B. 19). Constatou-se que as métricas  $X_2:X_{10}$ ,  $X_{12}$ ,  $X_{13}$ ,  $X_{18}:X_{21}$ ,  $X_{23}$  e  $X_{24}$ , em ambas as análises, foram consideradas como "*best fitted*" ( $p \text{ values} \geq 0,05$ ) e apresentaram magnitude "grande" ( $\eta^2 \geq 0,14$ ). Já as métricas  $X_{11}$ ,  $X_{19}$ , e  $X_{22}$  apresentaram uma magnitude "média", enquanto  $X_1$ ,  $X_{14}:X_{17}$  foram classificadas como "*no best fitted*", embora apresentassem magnitude "média".

Os resultados demonstraram um ARI de 0,27 (Subseção 3.5.9.1), ao comparar os agrupamentos derivados das análises HC e NHC da Abordagem 2. Esse valor sugere uma concordância moderada entre os dois métodos de agrupamento, indicando que, embora não sejam idênticos, há uma sobreposição significativa entre os *clusters* identificados por cada abordagem. A observação UTFPR foi excluída da análise, por se tratar de *outlier* uni e multivariada.

#### 4.6.1.1 Resultados da Análise HC (Hierárquica de Agrupamento ou Clustering)

##### 4.6.1.1.1 Abordagem 1

Todas as métricas ( $X_1$  a  $X_{24}$ ) foram levadas em consideração para a formação dos agrupamentos, sem avaliação individual de cada uma. Considerou-se que todas as métricas contribuem para o delineamento dos grupos.

Constatou-se a formação de dois (2) agrupamento distintos. O primeiro agrupamento, denominado “*cluster 1*”, é composto por 7 ICTs cujo valor médio do deste agrupamento foi de -0,32. O segundo agrupamento, “*cluster 2*”, conta com 33 ICTs, cujo valor médio deste agrupamento foi de 0,49.

A observação CPII foi excluída do *dataset* para evitar a condição de singularidade e permitir a continuidade da análise. A média geral da silhueta dos dois (2) agrupamentos encontrados foi de 0,35, sinalizando uma separação moderada entre eles.

##### 4.6.1.1.2 Abordagem 2

Considerou-se apenas as métricas que influenciaram diretamente a construção dos agrupamentos, após análise ANOVA *one way*. Constatou-se a formação de dois (2) agrupamentos distintos. O primeiro agrupamento, denominado “*cluster 1*”, é composto por 7 ICTs, cujo valor médio deste agrupamento foi de 0,46. O segundo agrupamento, “*cluster 2*”, conta com 33 ICTs, cujo valor médio deste agrupamento foi de 0,49. Não foi constatada nenhuma condição de singularidade. A média geral da silhueta dos dois (2) agrupamentos encontrados foi de 0,48, sinalizando uma separação moderada entre eles e indicando uma separação média entre os *cluster* maior que na Abordagem 1.

#### 4.6.1.2 Resultados da Análise NHC (Não Hierárquica Clustering)

##### 4.6.1.2.1 Abordagem 1

O critério de seleção de métricas seguiu o padrão da Abordagem 1 da análise HC. Constatou-se a formação de dois (2) agrupamentos distintos. O primeiro agrupamento, denominado “*cluster 1*”, é composto por 4 ICTs, cujo valor médio deste agrupamento foi de 0,88. O segundo agrupamento, “*cluster 2*”, conta com 37 ICTs, cujo valor médio deste

agrupamento foi de 0,49. Não foi constatada nenhuma condição de singularidade. A média geral da silhueta dos dois (2) agrupamentos encontrados foi de 0,35, sinalizando uma separação moderada entre eles.

#### 4.6.1.2.2 Abordagem 2

Considerou-se apenas as métricas que influenciaram diretamente a construção dos agrupamentos, após análise ANOVA *one way*. Constatou-se a formação de dois (2) agrupamentos distintos. O primeiro agrupamento, denominado “*cluster 1*” é composto por 7 ICTs, cujo valor médio deste agrupamento foi de 0,46. O segundo agrupamento, “*cluster 2*”, conta com 33 ICTs, cujo valor médio deste agrupamento foi de 0,49. Não foi constatada nenhuma condição de singularidade. A média geral da silhueta dos dois (2) agrupamentos encontrados foi de 0,48, sinalizando uma separação moderada entre eles e indicando uma separação média entre os *cluster* maior que na Abordagem 1.

### 4.6.2 Resultados da Análise da AFE por PCA

A análise AFE por PCA (Subseção 3.5.9.2, Apêndice C, Subseção C.2 e Apêndice B, Tabela B. 6) reduziu a dimensionalidade de um conjunto de dados, identificando padrões e relacionamentos entre as métricas ( $X_1:X_{24}$ ).

Foram utilizadas técnicas de Análise Multivariada para extrair os componentes principais latentes que explicassem a variabilidade observada nas métricas originais.

Foram identificadas as principais dimensões subjacentes que ajudam a explicar a estrutura de correlações entre as métricas.

Foram gerados *scores* que serviram para quantificar e representar a contribuição de cada variável latente (Componente Principal - CP<sub>i</sub>) para cada observação – ICTs – no conjunto de dados. Esses *scores* foram utilizados para calcular o impacto relativo de cada Componente Principal (CP<sub>i</sub>) em cada conjunto de dados.

Os *scores* foram multiplicados pelas suas variâncias correspondentes e, em seguida, padronizados para serem representados no intervalo entre [0,1], e ordenados para classificar ou comparar as observações analisadas – ICTs, com base nas características, subjacentes aos dados originais ( $X_1:X_{24}$ ) para tomada de decisões.

O Índice Integrado de Performance e Inovação em Pesquisa, Desenvolvimento Tecnológico e Prestação de Serviços Técnicos Especializados (IIP\_PD&S) foi elaborado

através de uma equação que considera a variância compartilhada dos componentes principais.

Nessas análises, os esforços foram centrados no resultado dos dados relativos ao intervalo temporal entre 2008 e 2018, para entender as tendências e padrões de desempenho da RFEPCT, ao longo desse período, e identificar quais ICTs se destacaram em diferentes momentos.

A observação UTFPR foi excluída da análise, por se tratar de *outlier* uni e multivariada.

#### 4.6.2.1 Testes Preliminares para a Análise Fatorial:

- Teste de Kaiser-Meyer-Olkin: foi considerado apropriado ( $KMO = 0,6$ ), indicando que é adequado prosseguir com as análises (indicador).
- Teste de Bartlett de Esfericidade: com  $p\text{-valor} < 0,001$ , rejeitou-se a hipótese nula de que as variáveis não estão correlacionadas nas populações (teste estatístico robusto).
- Medida de Adequação da Amostra Individual (MSAI): um total de 14 variáveis apresentaram valores maiores que 0,5. Essas variáveis são consideradas adequadas para a análise fatorial, a saber  $X_1:X_6$ ,  $X_8$ ,  $X_{10}:X_{13}$ ,  $X_{19}$ ,  $X_{20}$  e  $X_{23}$ . Já  $X_7$ ,  $X_9$ ,  $X_{14}:X_{18}$ ,  $X_{21}$ ,  $X_{22}$  e  $X_{24}$  apresentaram valores abaixo de 0,5, sendo consideradas menos adequadas para a análise fatorial exploratória, mas ainda podendo ser incluídas, considerados os objetivos específicos deste estudo (indicador).
- Matriz de Correlação: as correlações mais expressivas foram  $corr(X_3, X_4; r(41) = 0,78; p < 0,01)$ ,  $corr(X_3, X_5; r(41) = 0,93; p < 0,01)$ ,  $corr(X_3, X_6; r(41) = 0,90; p < 0,01)$ ,  $corr(X_4, X_5; r(41) = 0,78; p < 0,01)$ ,  $corr(X_4, X_6; r(41) = 0,72; p < 0,01)$ ,  $corr(X_5, X_6; r(41) = 0,83; p < 0,01)$ ,  $corr(X_3, X_{11}; r(41) = 0,46; p < 0,05)$ ,  $corr(X_4, X_{11}; r(41) = 0,65; p < 0,05)$ .
- Matriz de Correlação Anti-Imagem: os valores obtidos na diagonal (0,98) são maiores que 0,5, indicando viabilidade para continuar com a análise fatorial (indicador).
- Determinante da matriz de correlações: o valor foi de  $3 \times 10^{-3}$ , próximo a zero, indicando possíveis multicolinearidades entre as variáveis estudadas (indicador).

- Ajuste do modelo: apresentou um RMSR de 0,05 e um qui-quadrado empírico de 67,28,  $p < 0,01$ .
- Complexidade média do modelo: o valor foi de 1,6.

#### 4.6.2.2 *Extração de Componentes Principais:*

- Método de Extração: Componentes Principais (PCA).
- Número de componentes principais Extraídos: foram extraídos seis (6) Componentes Principais representativos, pois apresentaram autovalores maiores que um (1).
- Teste de hipóteses para Autovalores “*Eigenvalues*”: a extração dos seis Componentes Principais foi considerada suficiente.

#### 4.6.2.3 *Rotação de Componentes Principais:*

- Método de Rotação: *Varimax*.
- Variação Explicada após Rotação: os seis (6) Componentes Principais explicam 77,7% da variância total.

#### 4.6.2.4 *Definição dos Componentes Principais:*

- **Componente Principal 1:** intitulada “Canais de Pesquisa e Desenvolvimento Tradicionais”, explica 28% da variância total, acumulando uma soma dos quadrados (*SS loadings*) das cargas ( $0,09^2 + \dots + 0,11^2$ ) de 6,68. As métricas que mais contribuíram para este componente foram Outras Atividades ( $X_{13}$ ; 0,92), Projetos de P&D ( $X_5$ ; 0,91), Pesquisadores ( $X_2$ ; 0,90), Projetos de extensão ( $X_{12}$ ; 0,88), Projetos de pesquisa ( $X_4$ ; 0,83), Consultoria ( $X_{11}$ ; 0,80), Orientações de Pesquisa ( $X_3$ ; 0,78), Publicações Científicas ( $X_6$ ; 0,77) e Treinamentos ( $X_{10}$ ; 0,62).
- **Componente Principal 2:** intitulada “Canais de Cooperação em Inovação com SPP”, explica 15% da variância total e acumulada de 42%, acumulando uma soma dos quadrados cargas ( $0,00^2 + \dots + 0,88^2$ ) de 3,48. As métricas que mais contribuíram para este componente foram Desenhos Industriais com o SPP ( $X_{22}$ ; 0,99), Desenhos Industriais ( $X_{19}$ ; 0,90), Programas de computador com o SPP ( $X_{24}$ ; 0,88), Projetos de pesquisas em cooperação com SPP ( $X_8$ ; 0,58), Patentes Publicadas com o SPP ( $X_{23}$ ; 0,55).



- **Componente Principal 3:** intitulada “Canais de Iniciativas de Propriedade Intelectual e captação de Investimento”, explica 13% da variância total e acumulada de 56%, acumulando uma soma dos quadrados das cargas  $(0,73^2+\dots+0,26^2)$  de 3,24. As métricas que mais contribuíram para este componente foram Marcas ( $X_{18}$ ; 0,88), Programas de computador ( $X_{21}$ ; 0,73), Investimento Público em P&D ( $X_1$ ; 0,73) e Patentes Publicadas ( $X_{20}$ ; 0,65).
- **Componente Principal 4:** intitulada “Canais de Projetos Colaborativos Financiados pelos SPP”, explica 11% da variância total e acumulada de 66%, acumulando uma soma dos quadrados das cargas  $(0,15^2+\dots-0,01^2)$  de 2,55. As métricas que mais contribuíram para este componente foram Projetos de Extensão para SPP ( $X_{16}$ ; 0,85), Projetos de Pesquisas financiados pelo SPP ( $X_9$ ; 0,84), Orientação de Pesquisa financiada pelo SPP ( $X_7$ ; 0,65).
- **Componente Principal 5:** intitulada “Canais de Consultoria e Treinamento para SPP”, explica 6% da variância total e acumulada de 73%, acumulando uma soma dos quadrados das cargas  $(0,34^2+\dots-0,07^2)$  de 1,45. As métricas que mais contribuíram para este componente foram Consultoria para SPP ( $X_{15}$ ; 0,90) e Treinamentos para SPP ( $X_{14}$ ; 0,06).
- **Componente Principal 6:** intitulada “Canais de Atividades Diversificadas para SPP”, explica 5% da variância total e 77,7% acumulada, acumulando uma soma dos quadrados das  $(0,13^2+\dots-0,03^2)$  cargas de 1,24. A métrica que mais contribuiu para este componente foi Outras Atividades para SPP ( $X_{17}$ ; 0,72).

4.6.2.5 *Índice criado para classificar ou comparar as observações analisadas – ICTs – com base nas características subjacentes aos dados originais ( $X_1:X_{24}$ ) para tomada de decisões:*

- Índice: Intitulado “Índice Integrado de Performance e Inovação em Pesquisa, Desenvolvimento Tecnológico e Prestação de Serviços Técnicos Especializados”;
- Rótulo: IIP\_PD&S;
- Equação:  $IIP\_PD\&S = \sum_{i=1}^n (\% \text{variância do Fator}_i * \text{score do Fator}_i)$  ;
- Exemplo:  $IIP\_PD\&S = (0,28 * -0,61) + \dots + (0,05 * -0,66)$ ;

- Modelo de padronização implementado:  $IIP\_PD\&S_i = (X_{\max} - X_{\min}) / (X_i - X_{\min})$ , onde  $IIP\_PD\&S_i$  é o valor padronizado para o intervalo entre  $[0,1]$ .

4.6.2.6 *Desempenho das ICTs da RFEPCT ao Longo do Tempo, com base no índice criado, não pela ordem em cada item:*

- a. Mais Frequência nas Primeiras Posições: IFCE e IFES.
- b. Melhora Significativa: IFSP e IFSC.
- c. Desempenho Consistente: IFCE, IFES, CEFET/MG.
- d. Análise do Período Agregado (2008-2018): IFSC e IFES.
- e. Entradas entre as cinco (5) ICTs-MACE: IFSP e IFRS.
- f. Saídas entre as cinco (5) ICTs-MACE: IFPA e IFR.

## 5 Análise e Discussão

Em linhas gerais, a proposta do MVP consiste numa solução promissora que visa consolidar informações relevantes sobre a performance de instituições federais de ensino, auxiliando na tomada de decisões e na formulação de políticas públicas para aprimorar a qualidade das ações desenvolvidas. A ferramenta apresenta *design* amigável, boa navegabilidade e um conjunto abrangente de indicadores e métricas importantes. Contudo, após análise da seção 2, sugere-se melhorar a clareza e a definição dos indicadores, aprimorar a interface para facilitar a identificação dos dados e incluir uma maior variedade de instituições. Além disso, é recomendável criar um manual de utilização e inserir informações de referência das fontes e bancos de dados. A ferramenta possui potencial para ser customizada e expandida, tornando-se relevante para outras instituições e órgãos governamentais, como as Fundações de Amparo à Pesquisa (FAPs) dos estados.

A revisão de literatura realizada (Seção 2) abrange uma ampla gama de estudos e perspectivas sobre o ecossistema das IEPTC e os desafios enfrentados no campo da pesquisa, desenvolvimento e serviços. No entanto, apesar dos esforços e contribuições desses estudos, observou-se que até o momento nenhuma proposta da literatura analisada foi similar à solução tecnológica proposta nesta tese, através do MVP.

Os artigos revisados abordam diversas questões, como financiamento, recursos, articulação entre pesquisa e setores produtivos, infraestrutura e equipamentos, gestão, governança, capacitação e formação de recursos humanos, indicadores e métricas de avaliação, e internacionalização. No entanto, grande parte do referencial teórico analisado concentra-se em aspectos específicos ou aborda os desafios de maneira isolada, carecendo de soluções que tratassem do problema de forma longitudinal, envolvendo múltiplas fontes de dados e variáveis simultaneamente. A solução proposta buscou contribuir para a mitigação dessa lacuna.

Neste sentido, propõe uma solução tecnológica inovadora e abrangente, considerando todas as complexidades e desafios enfrentados pelas ICTs, ainda busca integrar os canais de PD&S para o fomento à intenção TH.

Em suma, a revisão de literatura ressaltou a originalidade e relevância da solução tecnológica proposta nesta tese, que tem potencial para transformar e aprimorar as práticas e políticas relacionadas à pesquisa, desenvolvimento e serviços no âmbito das ICTs. Essa mesma conclusão chegaram a grande maioria dos especialistas dos três setores

(ICTs, SPP e GOV) que avaliaram o MVP; também investigou em qual cenário nacional de CT&I brasileiro o ecossistema da RFEPC se desenvolveu durante sua primeira década de vigência, destacando a importância da colaboração entre ICTs, SPP e GOV para promover inovação e desenvolvimento apresentado na subseção 2.3 (Brasil, 2021b; 2022a; Etzkowitz *et al.*, 1995).

Nesse sentido, a evolução da teoria da Tripla Hélice foi revisada, mostrando como essa abordagem contribui para investigar o fenômeno apresentado na Subseção 2.4 (Etzkowitz *et al.*, 1995; 2000).

Na Subseção 2.5, o trabalho examinou a literatura sobre BI, DW, *dashboards* e KPIs aplicados em ecossistema das ICTs (Guitart *et al.*, 2015; Paz *et al.*, 2019). O estudo ressaltou o papel dessas tecnologias na melhoria da gestão de dados e informações, apoiando a tomada de decisões e o planejamento estratégico (Damyanov *et al.*, 2019). Além disso, casos práticos que ilustram a aplicação de BI, DW e *dashboards*, no contexto das ICTs, foram apresentados, como os estudos de Azevedo *et al.* (2021) no caso brasileiro e Bianchi *et al.* (2022) no caso internacional.

Apesar de o texto de Santos, Thiago Rizzi *et al.* (2021) não abordar diretamente a RFEPC e seus desafios na gestão e monitoramento da produção científica e tecnológica, ele destaca a importância dos sistemas de informação gerencial e a aplicação de tecnologias analíticas avançadas, como BI, *big data* e *Data Mining*, para apoiar a tomada de decisão em ICTs, para melhorar a gestão e o monitoramento de sua produção científica, tecnológica e prestação de serviços.

Nas próximas subseções, serão abordados alguns temas que contribuem para a compreensão e discussão das ICTs referentes as IEPTC no Brasil e no mundo.

Na sequência será promovida uma análise comparativa da educação profissional e tecnológica no Brasil e em outros países, focando nas semelhanças e diferenças em termos de organização, regulação e financiamento, ainda a implementação de tecnologias específicas em ecossistema das ICTs no contexto da Teoria TH, identificando desafios e oportunidades.

Posteriormente, será discutida a relevância e os desafios do uso de *dashboards* em ecossistema das ICTs, analisando como essas ferramentas podem impactar a gestão e a tomada de decisões, assim como o aprofundaremos desta análise, por meio de um estudo de caso. Por fim, será analisada e discutida a questão da pesquisa proposta relacionada à questão de pesquisa, o objetivo geral e à hipótese apresentada.

Essa jornada de investigação forneceu *insights* para a compreensão dos desafios e oportunidades enfrentados pelas ICTs no Brasil e no mundo, contribuindo para o desenvolvimento socioeconômico e inovativo de uma sociedade.

### **5.1 Análise Comparativa da Educação Profissional e Tecnológica no Brasil e em Outros Países: Semelhanças e Diferenças na Organização, Regulação e Financiamento**

A RFEPCT (Brasil, 2008), ao ser comparada com as redes de educação profissional e tecnológica de países como Alemanha, Canadá (Montreal), Chile, China e Finlândia, revela particularidades significativas. Enquanto esses países, diferentemente do Brasil, investem consideravelmente na educação profissional e tecnológica e passaram por uma mudança cultural que valoriza profundamente esta forma de educação, no Brasil, essa valorização e reconhecimento ainda parecem ser uma aspiração distante e, em muitos aspectos, utópica (Bank's, 2020; Brasil, 2016b).

A Subseção 4.3 destacou que, em comparação com outros países, o Brasil ainda encara desafios significativos para valorizar e investir adequadamente na educação profissional e tecnológica. Além disso, ressalta a urgência de estratégias eficazes e práticas ideais para enfrentar a complexidade da RFEPCT, que é marcada por uma diversidade de contextos em organização, regulação, financiamento e abrangência.

Neste contexto comparativo, realiza-se uma análise objetiva das (dis)similaridades entre a RFEPCT e as redes similares dos países considerados, utilizando a teoria TH como um marco teórico orientador dessa análise, cujo objetivo é colaborar para o desenvolvimento socioeconômico e inovativo da região na qual a instituição está inserida (Etzkowitz *et al.*, 1995; 2000). Os gestores, assim, podem propor ações e políticas baseadas em evidências, ao invés de em suas próprias convicções.

### **5.2 Análise da Implementação de Tecnologias Específicas em ecossistema das ICTs no Contexto da Teoria TH: Análise Crítica e Identificação de Desafios e Oportunidades**

A análise crítica dos resultados encontrados na implementação de tecnologias específicas, como BI, dashboard e DW, em ecossistema das ICTs, fundamentadas nas teorias de Análise de Negócio, Visualização de Dados e da Tripla Hélice de Inovação, constatou a necessidade de gerenciar e converter o capital passivo dos canais de PD&S

das ICTs no Brasil em ativos científicos, tecnológicos e de serviços (Perucchi *et al.*, 2015; Rossi, 2022).

Além disso, esse conjunto de tecnologias tem sido empregado para aprimorar a área de pesquisa em análise acadêmica (Paz *et al.*, 2019). Os resultados também sugerem a necessidade de explorar mais profundamente os Fatores e os Componentes Principais que influenciam o ecossistema de CT&I e PD&I (Analytics, 2018; 2019; Brasil, 2008; 2021b) brasileiro.

A identificação de barreiras e facilitadores para a adoção e implementação dessas tecnologias em ecossistema das ICTs, considerando fatores organizacionais, jurídicos, culturais e de infraestrutura tecnológica, são ainda uma realidade (De Negri, F. O. *et al.*, 2016; Niwa, 2014; Oliveira *et al.*, 2020; Peixoto *et al.*, 2021).

Esse ecossistema ainda carece de um amadurecimento cultural e sistêmico nesse sentido (Macedo, 2014; Santos *et al.*, 2015). A compreensão desses elementos pode auxiliar na elaboração de estratégias e políticas públicas voltadas ao fortalecimento delas, enquanto solução tecnológica viável.

A investigação ora realizada também destaca a necessidade de mais estudos que utilizem uma abordagem interdisciplinar e integrada aos canais de PD&S dos ecossistemas das ICTs, considerando suas especificidades e as necessidades dos diferentes atores envolvidos no processo de inovação (Azevedo *et al.*, 2021; Castillo *et al.*, 2018; Damyanov *et al.*, 2019; Liao *et al.*, 2019; Medina *et al.*, 2018; Wang *et al.*, 2013). Essa perspectiva pôde contribuir para a criação de ferramentas mais eficientes e adaptáveis, capazes de atender às demandas e desafios enfrentados por esse ecossistema, visto que poucos estudos têm abordaram essa possibilidade nos últimos anos.

Ao consultar a Seção 2, foi possível constatar que a falta de clareza e padronização quanto aos conceitos e definições, relacionados à inovação e propriedade intelectual, dificulta a adoção de um padrão comum para a comparação de resultados e a identificação de inovações efetivas em ecossistemas de ICTs (Bento, 2015; Lino, 2013; Marques *et al.*, 2022; Perucchi *et al.*, 2017; Rossi, 2022).

Ademais, os estudos existentes, geralmente, baseiam-se em análises bibliométricas e documentais, utilizando abordagens estatísticas simples e limitadas, sem explorar metodologias mais robustas que possibilitem um entendimento mais aprofundado dos resultados.

Os resultados, desta revisão de literatura, indicam diferenças metodológicas entre estudos anteriores, como a pesquisa de Chaves *et al.* (2017), e o MVP proposto,

destacando a necessidade de fomentar os canais PD&S da RFEPCT, em particular. Não obstante, a análise crítica da literatura (Subseção 2.4) evidenciou a crescente importância da aplicação de BI e suas tecnologias neste ecossistema, em geral. Estudos como os de Medina *et al.* (2018) e Damyanov *et al.* (2019) demonstram o impacto positivo que essas tecnologias podem ter na gestão de dados e informações, assim como no suporte à tomada de decisões e planejamento estratégico.

Outra observação importante é que, apesar de haver estudos que aplicam a BI e suas tecnologias relacionadas ao ecossistema das ICTs, muitos deles focam em métricas acadêmicas específicas ou em canais específicos, tais como patentes, projetos de pesquisa, publicações e grupos de pesquisa. Além disso, há uma ênfase maior em áreas acadêmicas como gestão, finanças e ensino, em detrimento da análise das interações ICTs, SPP e GOV (Subseção 2.1, 2.2 e 2.4).

Conforme analisado na Subseções 2.4, uma limitação identificada nos estudos relacionados à realidade do ecossistema da REFPCT, considerando sua singularidade (Subseção 2.1 e 4.2), é o foco em desenvolver soluções tecnológicas voltadas para atender às demandas de uma única ICTs, em detrimento de múltiplas instituições ou da própria REFPCT. Santos (2010) ressaltou que a utilização de fontes de dados para apoiar a tomada de decisão é mais comum, ao investigar ICTs federais brasileiras.

Embora haja limitações, alguns estudos oferecem contribuições, como o desenvolvimento de arquiteturas, protótipos de *dashboards* e modelos de projeto para a construção de *data warehousing* (Bassil, 2012; Castillo *et al.*, 2018). Guitart *et al.* (2015) e Hoffmann (2021) propuseram soluções voltadas para a gestão da atividade docente em ambientes virtuais de aprendizagem e nas ICTs federais brasileiras, focando nos canais de gestão administrativa.

No entanto, a maioria dos estudos analisados (Seção 2) tende a se concentrar em ecossistema interno das ICTs, limitando a visualização e organização dos dados para tomada de decisões endógenas. Bianchi *et al.* (2022) destacaram a lacuna na literatura quanto à abordagem dos canais de pesquisa e seu desenvolvimento no contexto das interações TH.

A plataforma aqui proposta pode ser considerada uma ferramenta importante para auxiliar na identificação e análise de variáveis relacionadas aos canais de PD&S em ecossistema das ICTs, permitindo maior transparência e possibilidade de auditoria nos resultados. Contudo, é fundamental atualizá-la continuamente e adaptá-la às mudanças e necessidades do setor. Além disso, ela facilita, por exemplo, a identificação de áreas de

pesquisa, colaboração entre docentes, ICTs e Redes de ICTs, além de possibilitar, interativamente, análises de correlações entre métricas disponíveis e filtros múltiplos, gerando no primeiro momento informações relevantes para tomada de decisão, pesquisa ou consulta pública. No entanto, a plataforma não substitui a necessidade de estudos aprofundados e a adoção de políticas públicas específicas para a promoção do ecossistema de CT&I e PD&I brasileiro.

Em resumo, este trabalho apresenta uma solução que aborda um problema real em ecossistema das ICTs, garantindo seu papel social, inovador e desenvolvedor regional, por meio de políticas, ações e atendimento às demandas dos diversos setores da sociedade relacionados aos desafios da gestão e ao monitoramento de sua produção e prestação de serviços, na área científica e/ou tecnológica.

A implementação de tecnologias como BI, *dashboards* e DW apresenta potencial para melhorar a gestão e a tomada de decisões nessas instituições. No entanto, é essencial considerar as limitações e lacunas identificadas na literatura e buscar soluções mais abrangentes e adaptáveis, que possam ser aplicadas a múltiplas ICTs e promovam o avanço do ecossistema de CT&I e PD&I no Brasil e no cenário internacional.

### **5.3 Análise da Relevância e Desafios do Uso de *Dashboards* em Ecossistema das ICTs**

A tecnologia de *dashboards*, surgida nos anos 1970, tem sido reconhecida como uma ferramenta valiosa para a análise acadêmica e visualização de dados em ecossistema das ICTs (Mould *et al.*, 2016; Stewart *et al.*, 2022; Wang *et al.*, 2015). Entretanto, é importante considerar suas limitações e possibilidades no contexto mais amplo da TH. A aplicação de *dashboards* pode ser estendida aos sistemas de informação gerencial e de suporte executivo, que buscam otimizar o tempo de acesso às informações (Stair *et al.* (2010).

A crescente importância da análise acadêmica (Paz *et al.*, 2019) torna o uso de *dashboards* essencial ao ecossistema das ICTs. De acordo com os estudos analisados, durante a revisão da literatura, a popularidade do *dashboard* se deve à sua capacidade de ajudar a monitorar KPIs e identificar tendências para melhorar o desempenho das instituições (Rabova *et al.*, 2014). Contudo, foi observada a carência de soluções tecnológicas específicas para auxiliar tomadores de decisão no planejamento e na elaboração de políticas baseadas em dados relacionados aos canais de PD&S, embora já



existam algumas soluções tecnológicas desenvolvidas para atender às demandas de outros canais do ecossistema das ICTs, como os setores administrativo e educacional (Seção 2 e Apêndice B, Quadro B. 1 e Quadro B. 17).

A importância do *UX design* é enfatizada em alguns estudos, especialmente na criação de *dashboards* baseados em BI (Erdos *et al.*, 2019), pois busca entender as necessidades e expectativas dos usuários e desenvolver soluções visuais que ofereçam uma experiência intuitiva e atraente. O *dashboard* é um exemplo de aplicação do *UX design* em ferramentas eficientes e adaptáveis (Martins *et al.*, 2022). Entretanto, Erdos *et al.* (2019) destacaram a necessidade de considerar os custos e benefícios financeiros a longo prazo, ao tomar decisões de investimento em *software*. Por esse e outros motivos, o *UX design* implementado foi planejado para ser de fácil manutenção e adaptabilidade, uma vez que essa plataforma ainda está no estágio de um MVP.

Os KPIs são valores quantificáveis e adotados para definir, medir e monitorar o desempenho das instituições ao longo do tempo, em relação aos seus objetivos organizacionais (Albats *et al.*, 2018; Rabova *et al.*, 2014). Eles podem ser baseados em tempo, custo e qualidade (Bertagnolli, 2022), servindo como guias para acompanhar metas e estratégias (ABNT, 2000). A implementação de KPIs deve ser cuidadosamente planejada e alinhada com os objetivos estratégicos das instituições envolvidas.

A revisão da literatura não apenas constatou a relevância do uso de BI, DW, *dashboards* e KPIs em ecossistema das ICTs, como também evidenciou lacunas na literatura relacionadas à aplicação de *dashboards* e DW associados a BI para propor soluções tecnológicas no contexto da TH. Desafios como integração de múltiplas fontes, padronização de processos e estudos longitudinais que envolvam grupos de instituições atuando em rede ainda precisam ser superados.

Por fim, é essencial que se invista em capacitação e formação de gestores e profissionais envolvidos e/ou inseridos em ecossistema das ICTs, de modo a garantir o uso adequado e eficiente das tecnologias de BI, *dashboards* e DW. A adoção dessas tecnologias, por si só, não garante a melhoria da gestão de dados e informações, sendo necessário que os usuários estejam aptos a interpretar e utilizar as informações e análises geradas de forma estratégica e assertiva.

Outro ponto a destacar é a falta da cultura de dados dos gestores brasileiros, os quais, em parte, tendem a tomar decisões estratégicas baseadas em convicções e não em dados que atendam aos pressupostos da Lei de Segurança da Informação (Brasil, 2018a).

Em síntese, a análise realizada destaca a necessidade de avanços na pesquisa e prática para superar as limitações identificadas e ampliar o escopo de aplicação dessas tecnologias. Ao considerar esses aspectos, é possível alavancar o potencial das ferramentas de BI para melhorar a gestão de dados e informações e apoiar a tomada de decisões e o planejamento estratégico em ecossistema das ICTs.

#### **5.4 Análise da Questão de Pesquisa e das Hipóteses Apresentadas**

A análise das avaliações e resultados apresentados pelos especialistas (Subseção 4.3 e 4.4 e Apêndice B, Quadro B. 3) concentrou-se em validar ou refutar a hipótese inicial, assim como também a questão da pesquisa apresentada neste estudo.

Nove especialistas, representando os três setores da TH, foram convidados para avaliar a proposta tecnológica detalhada neste estudo. O objetivo principal foi aprimorar a capacidade de resposta dos gestores das ICTs às demandas institucionais, produtivas e governamentais, incluindo aquelas provenientes dos órgãos de controle governamentais e, indiretamente, reduzir o hiato entre a teoria consolidada e a pesquisa empírica atual, sem ignorar a heterogeneidades intrínseca do SNI (De Negri, F. O. *et al.*, 2016), a participação e a contribuição das ICTs, nesse ecossistema.

Mesmo integrando uma rede como a RFEPCT, é essencial considerar as individualidades e heterogeneidades presentes, característica essa presente em outras ICTs nacionais ou internacionais (Andrade *et al.*, 2022; Berbegal-Mirabent *et al.*, 2019; Cesaroni *et al.*, 2015; Fan *et al.*, 2015). Isso se reflete, não apenas na constituição dessas entidades, mas, também, na maneira como interagem com os canais de PD&S, influenciando diretamente os resultados alcançados (De Negri, F. O. *et al.*, 2016). Isso posto, é possível afirmar de que a RFEPCT era heterogênea até 2018. A seguir, serão apresentados os resultados alcançados, focando nos atributos e potencialidades da solução propostas.

Os especialistas foram unânimes em destacar a relevância e a efetividade da solução tecnológica aqui proposta. Esse consenso entre os nove especialistas reflete uma perspectiva promissora para sua implementação futura e já aplicabilidade na versão MVP, demonstrando ser uma ferramenta essencial para decisões estratégicas no setor governamental e produtivo, em especial, em auxiliar os gestores no atendimento às exigências legais e administrativas dos órgãos de controle governamentais, bem como das

demandas por soluções solicitadas pelos SPP, contribuindo para o desenvolvimento regional.

Ao se analisar individualmente as considerações dos especialistas, observou-se que a customização da ferramenta para outros ecossistemas brasileiros foi um ponto abordado, destacando-se a potencialidade de expansão para futuros clientes (Especialista 2). Além disso, salientou-se a necessidade de melhorias em características específicas, incluindo a definição de outros indicadores (Especialista 3, 5 e 8).

No que concerne à usabilidade da ferramenta, a implementação de um manual de utilização e o ajuste na interface foram propostos como estratégias para tornar a solução mais acessível e funcional (Especialista 5 e 6). Foi ressaltado, ainda, o potencial da ferramenta para a gestão otimizada de políticas, indicando-se a necessidade de otimizações que permitam uma inserção mais harmoniosa de novas variáveis e parâmetros (Especialista 7 e 9).

Ressalta-se que tais necessidades podem ser facilmente atendidas desde, de que haja as condições financeiras e infraestrutura para a implementação das melhorias e implementação do MVP como uma ferramenta de produção. Por ora, é possível comprovar sua utilidade e inovação, enquanto uma solução tecnológica no contexto ao qual ela foi proposta para ser validada.

Com base no *feedback* dos especialistas e na análise do contexto apresentado, é possível inferir que há um suporte substancial para a aceitar a hipótese proposta. O consenso entre os especialistas reforça a noção de que o MVP proposto é, de fato, relevante e inovador, proporcionando um meio eficaz de apoiar a tomada de decisão, análise, gestão e monitoramento.

Além disso, a menção à capacidade da ferramenta de ajudar a atender as exigências legais e administrativas do GOV e, também, as demandas por soluções dos SPP, foi corroborada pelos comentários dos especialistas sobre o potencial da ferramenta para aprimorar a tomada de decisões estratégicas e facilitar a gestão otimizada dos referidos canais. Assim, o MVP proposto também está alinhado com as teorias de Análise de Negócios, Visualização de Dados e da Tripla Hélice de Inovação, podendo proporcionar uma integração significativa e otimizada da gestão dos dados dos canais de PD&S em ICTs, desencadeando uma série de benefícios substanciais para todas as partes envolvidas, conforme a avaliação dos especialistas.

A partir dessas avaliações sugere-se como possíveis ações e ou diretrizes para mitigar o problema identificado neste trabalho:

## **I. Integração e Análise de Negócios**

- Tomada de Decisão Aprimorada: com base em uma análise meticulosa dos dados, os gestores podem receber *insights* profundos para guiar suas estratégias, atendendo mais eficientemente às demandas variadas e complexas que surgem tanto no ambiente institucional quanto governamental.
- Integração de Dados: a unificação dos dados de diversos setores facilita a identificação de oportunidades valiosas e desafios emergentes, permitindo uma resposta mais célere e alinhada às exigências apresentadas.

## **II. Visualização de Dados**

- Facilitação do Entendimento: por meio da representação gráfica e interativa das informações, é possível simplificar a interpretação dos dados, o que confere uma maior fluidez e eficácia no processo de tomada de decisões.
- Rápida Identificação de Tendências e Padrões: a visualização facilitada permite que tendências e padrões sejam percebidos mais rapidamente, oferecendo um fundamento sólido para estratégias bem-informadas e decisões produtivas.

## **III. Tripla Hélice de Inovação**

- Colaboração Multi-setorial: a interação promovida entre universidades, indústrias e governo engendra um ambiente favorável à inovação e ao crescimento regional, através da construção de uma rede de cooperação robusta e resiliente.
- Resposta Rápida às Demandas Governamentais: o entendimento facilitado entre os três braços da TH possibilita uma resposta mais ágil e eficiente às demandas governamentais, alinhando os objetivos e facilitando o cumprimento das exigências regulatórias.

## **IV. Benefícios Adicionais**

- Aprimoramento da Usabilidade e Acessibilidade: a personalização da ferramenta de BI para atender às especificidades de diferentes

ecossistemas, promove uma maior acessibilidade, favorecendo a sua integração em variados ambientes e atendendo, de maneira satisfatória, às demandas dos SPP.

- **Potencial para Expansão:** a solução, por ser expansível, permite a incorporação harmoniosa de novas variáveis e parâmetros no sistema, facilitando sua implementação em outros contextos brasileiros e contribuindo para um desenvolvimento regional sustentável e integrado.

A implementação da solução tecnológica em BI proposta está fundamentada em teorias e práticas estabelecidas, servindo como uma ferramenta promissora para a gestão de dados integrada e otimizada, essencial para atender às variadas exigências e demandas institucionais, produtivas e governamentais, promovendo assim o desenvolvimento regional, através de decisões estratégicas e bem fundamentadas. Essa perspectiva é corroborada pela análise das avaliações dos especialistas, que validaram sua eficácia e ressaltaram sua relevância e utilidade pública, indicando um cenário positivo para sua implementação, desde que se realizem os ajustes e melhorias sugeridos, nas áreas identificadas como prioritárias.

### **5.5 Análise dos Resultados Descritivos da Rede Federal (2008-2018), da Revisão da Literatura e da Mineração de Dados**

Conforme os resultados apresentados nas Subseções 4.5 e 4.5, foi possível identificar pontos cruciais relacionados aos resultados dos canais de PD&S na RFEPCT.

A Figura A. 1 (Apêndice A), apresenta a distribuição geográfica das ICTs que compõem a RFEPCT. É importante considerar a expansão histórica desse conjunto, que tem suas raízes na primeira iniciativa de educação profissional no Brasil, estabelecida em 1686 (Apêndice A).

Ao analisar os resultados dos estudos supracitados, percebe-se uma concordância significativa sobre a influência histórica e geográfica na atual configuração das ICTs nos SNI. Observa-se que a colaboração TH, em especial as ICTs, tende a apresentar resultados mais expressivos em regiões em desenvolvimento. Especificamente, em países como Brasil e Espanha, isso é mais notável nas áreas litorâneas. Entretanto, na China, a colaboração TH também se destaca em regiões em desenvolvimento, mas não necessariamente nas áreas litorâneas.

A preocupação com a metodologia adequada para análise desses sistemas complexos também é um ponto de convergência, com (Berbegal-Mirabent *et al.*, 2019) identificando uma lacuna entre a teoria e a pesquisa empírica e (Leydesdorff *et al.*, 2019) e (Andrade *et al.*, 2022) advogando por uma análise que preserve a complexidade dos sistemas.

Este estudo buscou colaborar e contribuir com uma proposição tecnológica que diminua a lacuna entre a teoria e a pesquisa empírica, preservando a complexidade do SNI e as heterogeneidades das ICTs, anda que pertençam a uma rede, como a RFEPCT.

Ao analisar criticamente os estudos de Andrade *et al.* (2022), Li *et al.* (2019) , Leydesdorff *et al.* (2019) e Berbegal-Mirabent *et al.* (2019), contatou-se divergências na abordagem, especialmente na avaliação dos resultados de políticas de descentralização e regionalização. Enquanto aqui é apresentado um quadro de crescimento da expansão e cobertura da RFEPCT no Brasil – até por se tratar de um conjunto de ICTs ainda “novo”, Leydesdorff *et al.* (2019) delineou um cenário menos promissor para a Espanha, apontando para a centralização da inovação em poucas metrópoles e considerando insatisfatórios os resultados das políticas de regionalização implementadas após 1978.

Com relação, ainda, aos *investimentos* ( $X_1$ ) entre as ICTs, foi possível constatar que sua variação foi de 136,6%, isso com alguns recebendo muito mais ou muito menos do que a média. No entanto, os resultados da correlação entre as nove (9) variáveis relacionadas a interação com os SPP e os *investimentos* ( $X_1$ ) são mistos. Enquanto duas (2) variáveis mostraram uma correlação moderada, a maioria não apresentou nenhuma correlação significativa. Isso sugere que os  $X_1$  não necessariamente se traduzem em maior interação com os SPP de forma linear; aspecto esse que foi também abordado por Musa *et al.* (2019).

Esse é um indicador de que  $X_1$  não foi uniformemente distribuído entre as ICTs durante a 1ª década. A UTFPR destacou-se com o maior investimento, enquanto o CPIO não registrou nenhum investimento. Salienta-se que o dado foi extraído do Censo da Educação Superior e o registro oficial era zero unidade monetária e não nulo ou ausente. Manteve-se o resultado conforme registrado. Entretanto, contatos foram realizados via Sistema de Informação ao Consumidor. A resposta fornecida, porém, não permitiu realizar possível correção.

Para relatar os dados discretos, optou-se por utilizar a mediana e a moda, consideradas abordagens apropriadas para descrever a tendência central desses dados, tal como o número *de pesquisadores* ( $X_2$ ). A escolha da mediana justifica-se por sua robustez

como medida, não sendo afetada por valores extremos, enquanto a moda foi empregada para indicar o valor que ocorre com maior frequência nos dados discretos ( $X_2$ ).

O Coeficiente de Variação (CV) é uma medida estatística de dispersão relativa que permite comparar a variabilidade entre séries de dados diferentes, independentemente das unidades de medida. Ele é expresso como a razão entre o desvio padrão e a média, multiplicada por 100, para transformar o resultado em porcentagem. CV acima de 30% sugere que os dados apresentam alta variabilidade em relação à média dos dados (Brown, 1998). Destaca-se que dados normalmente distribuídos tendem a apresentar um CV menor do que dados que não são. Os dados discretos e contínuos deste trabalho não apresentaram distribuição normal, embora representem os resultados do total de 41.360 amostras.

Todas as 24 variáveis numéricas ( $X_1:X_{24}$ ) apresentaram coeficientes de variação após a 1ª década de funcionamento acima do valor sugerido. Estudos posteriores devem considerar esses resultados para efeito de comparação de resultados.

Após 10 anos, foi possível identificar, por exemplo, a importância que as RFEPCT, por meio de suas ICTs prestaram em relação ao canal “*consultorias*” ( $X_{11}$ ), registrado no total 4.145 serviços de consultoria nos mais diversos campos do conhecimento e setores, em especial no GOV. No entanto, apenas uma fração mínima dessas consultorias foi direcionada aos SSP registrado no total 15, destacando uma possível lacuna na colaboração entre ICTs-SPP. Esse é um aspecto que foi sublinhado por Perucchi *et al.* (2015), ao relatar que dentre os pesquisadores pesquisados e que informaram algum tipo de prestação de serviço, não foram encontrados registros referentes à consultoria. Já D'este *et al.* (2007) destacaram a negligência em relação a questões importantes, como projetos conjuntos de pesquisa, consultoria e treinamento, apesar da variedade de canais relacionados à transferência do conhecimento.

Em relação às publicações científicas, este estudo constatou que uma proporção significativa de publicações foi em coautoria entre as ICTs. Isso indica uma colaboração saudável entre diferentes instituições dentro da RFEPCT, entretanto, pode ser considerada uma interação ainda endógena do conhecimento. No entanto, ainda há margem para melhorias, como enfatizado por Bianchi *et al.* (2022).

O investimento em professores com Dedicção Exclusiva (DE) demonstra ser fundamental para o desenvolvimento nacional em CT&I e PD&I. Os dados analisados indicam que 87,41% dos *projetos de pesquisa* ( $X_4$ ) foram submetidos por *pesquisadores* ( $X_2$ ) com DE. Tal constatação não apenas ressalta a intensa atividade de pesquisa desses

docentes, mas também reforça a tese de que a atuação dos professores com DE, seja na pesquisa ou na extensão, se estende significativamente além das horas formalmente registradas, especialmente na carreira de Ensino Básico, Técnico e Tecnológico (EBTT) vinculada à RFEPCT.

Apenas a Universidade Tecnológica Federal do Paraná (UTFPR) é regida pelo magistério superior, diferenciando-se nesse aspecto. A proporção substancial de docentes que possuem dedicação exclusiva destaca seu compromisso não apenas com a pesquisa, mas também com o ensino, evidenciando a importância dessa categoria profissional para a qualidade da educação e para o avanço da pesquisa no país. Essa observação reforça a argumentação de que um investimento contínuo e sólido nos professores com DE, principalmente naqueles vinculados à RFEPCT, é uma estratégia eficaz e necessária para fomentar o desenvolvimento sustentável e inovador do Brasil. Esse resultado é consistente com a literatura consultada. Wang *et al.* (2013) argumentaram que a correlação reforça a ideia de que uma produção científica robusta pode levar a inovações patenteadas. Essa relação é corroborada pelos dados da RFEPCT. Eles revelam uma correlação muito forte ( $r=0.873$ ) entre essas duas variáveis. Isso sugere que há uma associação robusta entre a atividade de pesquisa acadêmica, conforme medida pelas publicações, e a inovação, conforme medida pelas patentes.

Ressalta-se, no entanto, que essa correlação significativa e forte não implica causalidade. É possível que outros Componentes Principais, como financiamento de pesquisa, políticas institucionais ou cultura de inovação, também desempenhem um papel importante nessa relação. Além disso, comparado com estudos anteriores, a correlação encontrada neste estudo é relativamente alta, o que pode indicar uma tendência crescente de alinhamento entre pesquisa acadêmica e atividades de inovação nas instituições analisadas.

A variação de *pesquisadores* ( $X_2$ ) sugere que algumas instituições têm equipes de pesquisa maiores ou mais ativas, o que pode indicar um foco mais robusto na pesquisa, por exemplo.

Os resultados de *orientações de pesquisa* ( $X_3$ ) e projetos de pesquisa ( $X_4$ ) sugerem uma grande variação no número de orientações e projetos de pesquisa entre diferentes instituições. A discrepância significativa entre a mediana e a moda para ambas as variáveis indica que existem algumas instituições com números muito elevados de orientações e projetos de pesquisa, o que pode estar distorcendo a distribuição dos dados. Isso sugere que, enquanto algumas instituições estão muito envolvidas em atividades de



pesquisa, outras têm envolvimento significativamente menor. Essas diferenças podem ocorrer devido a vários fatores, incluindo recursos, foco institucional e capacidade de pessoal.

Os valores elevados para  $X_3$  e  $X_4$  destacam o compromisso das ICTs para o desenvolvimento acadêmico; neste contexto elas podem ser consideradas métricas vitais para o desenvolvimento acadêmico. Os dados fornecidos mostram uma matriz de correlação entre essas duas variáveis. A correlação de *Pearson* entre  $X_3$  e  $X_4$  foi ( $r = 0,83$ ,  $p < 0,01$ ), indicando uma correlação positiva e significativa.

É importante ressaltar que correlação não implica causalidade, mas, no contexto das ICTs, faz sentido que um maior engajamento em orientações de pesquisa esteja associado a um maior número de projetos de pesquisa, uma vez que ambos são indicativos de uma cultura acadêmica ativa e produtiva. Além disso, a significância estatística ( $p < 0,01$ ) desta correlação fortalece o argumento de que há uma associação substantiva entre essas duas variáveis, e não é uma relação espúria. Portanto, pode-se inferir que ICTs que registram altos números, tanto em orientações quanto em projetos de pesquisa, demonstram um comprometimento significativo para o desenvolvimento acadêmico, sendo provável que estejam mais envolvidas em atividades de pesquisa e orientação. Essa correlação robusta serve como uma evidência baseada em dados de um compromisso focado na pesquisa e no desenvolvimento acadêmico, o que pode ser vital para fomentar uma cultura de excelência e inovação em pesquisa

*Publicações* ( $X_6$ ) é considerado um indicador de produtividade acadêmica. A UTFPR lidera em termos de  $X_6$ , o que pode refletir a qualidade da pesquisa ou o foco na divulgação acadêmica.

Em relação às *Patentes*, Suzuki (2018) argumenta que, embora elas sejam um indicador de inovação comumente utilizado, sua eficácia como métrica é questionável; uma conclusão corroborada neste trabalho. Destarte, é possível afirmar que, em uma escala ordinal de importância e valor inovativo, uma patente depositada é muito menos valorativa e importante do que uma concedida<sup>22</sup>, licenciada<sup>23</sup> ou cedida<sup>24</sup>. A liderança de

---

<sup>22</sup> A *Concessão* é o ato pelo qual o órgão responsável reconhece oficialmente uma invenção como patente e confere direitos exclusivos ao inventor ou titular.

<sup>23</sup> O *Licenciamento* é um acordo pelo qual o titular da patente permite que outra parte use sua invenção, geralmente em troca de *royalties* ou outras compensações, sem transferir a propriedade dos direitos da patente.

<sup>24</sup> A *Cessão* é o ato de transferir esses direitos de uma pessoa ou entidade para outra.

CEFET/MG e IFFAR em  $X_{20}$  sugere um foco significativo em pesquisa aplicada e inovação, com potencial de evoluir para concedida, licenciada ou cedida.

Uma observação preocupante é que nove (9) ICTs, até o período analisado, não haviam registrado nenhuma atividade relacionada a patentes publicadas, por exemplo. Isso destaca uma possível desconexão entre a pesquisa realizada e sua aplicação prática ou possível comercialização ou interação com os SPP.

A análise longitudinal de uma década de dados das ICTs que compõem a RFEPCT revela diversidade nos investimentos, produção acadêmica e inovação entre as instituições. Algumas, como a Universidade Tecnológica Federal do Paraná (UTFPR) e o Instituto Federal de São Paulo (IFSP), destacam-se em múltiplas categorias, constituindo-se como possíveis referências para outras instituições da Rede. Nesse contexto, evidencia-se a importância de reconhecer e apoiar instituições com potencial de crescimento em áreas específicas. Sendo a primeira considerada um *outlier* da RFEPCT.

A análise dos dados indica diferenças significativas nos investimentos entre as instituições, *insight* que se revela fundamental para compreender as prioridades de financiamento e para identificar setores que possam demandar a alocação de mais recursos. Essas observações são valiosas para a tomada de decisões estratégicas, especialmente no que diz respeito à alocação de recursos para impulsionar a excelência em pesquisa e inovação na RFEPCT.

Por fim, o estudo identificou que variáveis com medianas mais baixas, como o registro de *programas de computador* ( $X_{21}$ ), representam áreas com potencial de crescimento. Assim, sugere-se investir em treinamento e recursos nesses segmentos, com o intuito de beneficiar a RFEPCT de maneira ampla e potencializar sua interação TH.

Identificar as ICTs que se destacam em áreas específicas pode elucidar as especialidades de cada uma, potencializando a replicação de práticas bem-sucedidas em diferentes contextos institucionais. Contudo, apesar dos sinais positivos observados nos canais de PD&S da RFEPCT, existem canais que necessitam de atenção e aprimoramento. Entre elas, destacam-se a colaboração com os setores produtivos, a comercialização de pesquisas e a otimização dos investimentos públicos. Esses aspectos, quando adequadamente planejados e monitorados, têm o potencial de fortalecer significativamente as ICTs, conforme ressaltado neste estudo.

## 5.6 Análise conjuntural da mineração de dados aplicada à Rede Federal (2008-2018)

A análise integrada dos dados (Apêndice B, Tabela B. 7) através dos métodos HC, NHC (*K*-médias ou *K-means*), considerando os agrupamentos formados pela Abordagem 2, em conjunto com a classificação realizada pelo índice IIP\_PD&S e dos seis (6) Componentes Principais criados, fundamentados na técnica de análise fatorial PCA, permitiu uma compreensão profunda e detalhada do resultado apurado da RFEPCT, entre 2008 e 2018 (Apêndice B, Figura B. 18: Estudo comparativo de Agrupamento Hierárquico das Abordagens 1e 2 para o intervalo temporal entre 2008 e 2018 da RFEPCT.Figura B. 18 e Figura B. 19). Essa análise possibilitou constatar sua taxionomia e identificar os subconjuntos naturais formados pelas ICTs que as constituem, assim como os padrões e relações existentes entre seus canais de PD&S e sua contribuição ao modelo TH de inovação.

O “*cluster 1*”, em qualquer modelo analítico utilizado, agrupa as instituições com os indicadores mais elevados de IIP\_PD&S, indicando que este subconjunto natural formado por essas ICTs, ao longo do tempo, demonstrou melhores desempenhos com base nos seis (6) Componentes Principais criados. As instituições em destaque são: IFES (2), IFCE (3), IFSP (4), CEFET/MG (5), IFRS (6) e IFRN (7).

Destaca-se que, embora, tenha sido constatadas divergências entre os agrupamentos formados pela análise hierárquica e não hierárquica, esse resultado, em si, não é e nem deve ser considerado um problema, de acordo com a literatura especializada da área (Fávero *et al.*, 2020; Field, 2017; Hair *et al.*, 2019). A recomendação recorrente pelos especialistas é de que, ao invés de visualizar as divergências como obstáculos, deve-se encará-las como fontes ricas de *insights* adicionais, que podem proporcionar uma compreensão mais profunda e detalhada dos dados em estudo, facilitando a extração de conclusões mais robustas e bem fundamentadas.

Ao analisar a coerência entre os agrupamentos hierárquicos e não hierárquicos, foi possível constatar que, em geral, seis ICTs mantiveram-se no agrupamento em que se concentram os melhores resultados, conforme já apresentado acima, com exceção do IFSC. Tal divergência sugere que a migração do IFSC presente no “*cluster 1*” (HC) para o “*cluster 2*” (NHC), pode indicar que essa ICTs possui uma característica distinta que é mais evidenciada, dependendo da abordagem de análise utilizada. É possível que a forte presença no Componente Principal 2, que representa os “Canais de Cooperação em

Inovação com SPP”, seja um diferencial significativo para o IFSC, indicando uma estratégia de cooperação bem-sucedida com os SPP.

Ao analisar a distribuição geográfica dessas instituições, verifica-se que três (3) delas estão situadas na região Sudeste (IFES, IFSP e CEFET/MG), duas no Nordeste (IFCE e IFRN) e uma no Sul (IFRS). Isso demonstra uma predominância de ICTs de alto desempenho na região Sudeste, consequência da maior concentração de recursos e expertise técnica nessa região. Entretanto, para compreender completamente a dinâmica que orientou a constituição natural deste “*cluster 1*”, é fundamental detalhar os Componentes Principais exógenos e endógenos que, possivelmente, contribuíram para essa formação. Destaque-se que o IFSC está localizado na região Sul, juntamente com o IFRS, compartilhando assim as mesmas características e oportunidades regionais. Tal constatação sugere que a Rede tende a se constituir em uma formação similar à de Espanha, situação não desejada para países que planejam o desenvolvimento de PD&S de forma integrada.

A região Sudeste, especialmente em Minas Gerais, destaca-se no Componente Principal 1 “Setor de Pesquisa e Desenvolvimento Tradicionais”. Instituições como IFSP e IFES ressaltam-se também em iniciativas de propriedade intelectual, contribuindo assim para uma forte presença no “*cluster 1*”. Já no Sul, percebe-se a cooperação destacada em inovação com SPP, especialmente por parte do IFSC, aliado a uma performance equilibrada do IFRS e IFRN em diversos canais de inovação. Essa região evidencia uma capacidade institucional versátil, demonstrada pela distribuição equilibrada de instituições nos agrupamentos hierárquicos e uma tendência de bom desempenho a partir do IIP\_PD&S.

O Nordeste não fica atrás, com instituições como IFCE e IFPE sobressaindo-se em projetos colaborativos e consultoria para SPP, respectivamente. A região mostra força com representantes no “*cluster 1*” marcando presença no agrupamento hierárquico.

Ademais, observou-se que o Centro-Oeste abriga predominantemente instituições no agrupamento hierárquico, ainda que apresente variabilidade no índice IIP\_PD&S. O IFMA distingue-se, alcançando também o agrupamento não hierárquico. Em contrapartida, a região Norte aponta para uma oportunidade de crescimento, caracterizando-se por abrigar instituições de hierarquia média, mas com perspectivas de melhoria, como demonstrado pelo desempenho elevado do IFAM, com base na classificação pelo índice proposto IIP\_PD&S.

Por fim, até o período abordado neste estudo, observou-se dez (10) ICTs-MACE, com destaque de três ICTs: duas da região Nordeste (IFPB e IFPE) e uma da região Norte (IFPA). Esse dado evidencia o progresso e eficácia das políticas públicas e governamentais de estímulo à interação e colaboração entre ICTs-GOV-SPP. O objetivo dessas políticas é impulsionar o desenvolvimento socioeconômico por meio da inovação, conforme demonstrado pelos resultados alcançados nos canais de PD&S nessas regiões. Por enquanto, ainda não foi identificado qualquer ICTs da Região Norte e Centro-Oeste no “*cluster 1*” e sete (7) ICTs-MACE.

Assim, o panorama nacional revela uma ampla e diversificada distribuição de ICTs, com cada região apresentando peculiaridades e potenciais distintos para desenvolvimento no ecossistema de STI.

O Sistema Nacional Brasileiro de Ciência, Tecnologia e Inovação é composto por órgãos governamentais, instituições de ciência e tecnologia (por exemplo, as instituições de ensino e pesquisa e fundações de pesquisa), empresas e outras entidades. Sua função é promover a inovação e o desenvolvimento tecnológico no Brasil.

A presença dominante das instituições do Sudeste na supracitada ordenação pode ser um reflexo do forte ecossistema de inovação dessa região, com a presença de parques tecnológicos, incubadoras e empresas de base tecnológica.

Para compreender integralmente a dinâmica que guiou a formação natural desse “*cluster 1*”, torna-se essencial investigar os fatores exógenos e endógenos envolvidos, a partir de dados como os fornecidos pelo Instituto Brasileiro de Geografia e Estatística, através do Censo Demográfico, da Pesquisa Industrial Anual e da Pesquisa de Inovação Tecnológica.

Entre 2008 e 2018, o Brasil viu um aumento na pesquisa e inovação, especialmente no setor industrial. A Pesquisa Industrial Anual e a Pesquisa de Inovação Tecnológica, realizadas pelo IBGE, mostram que as regiões Sudeste e Sul são os principais polos industriais e tecnológicos do Brasil.

Essa região contém as dez (10) cidades com maior Índice de Desenvolvimento Humano Municipal (IDHM) do país. Este resultado é impulsionado por fatores como a presença de metrópoles significativas, centros acadêmicos e de pesquisa renomados, infraestrutura logística ampla e um vasto mercado consumidor. São Paulo e Rio de Janeiro destacam-se como epicentros de indústrias de grande porte e alta tecnologia, sendo responsáveis pela maior parcela do Produto Interno Bruto nacional e concentrando mais da metade da população brasileira.

O setor industrial na região supracitada é diversificado, abrangendo áreas como eletrônica, automobilística, química e metalurgia, com São Paulo e Minas Gerais como principais *hubs*. A região ainda se caracteriza como o principal polo de inovação e tecnologia do Brasil, hospedando destacados centros universitários e de pesquisa, o que é contatado, ao se analisar entre as sete (7), baseadas no índice proposto neste estudo.

A região Sul, especialmente os estados de Santa Catarina e Rio Grande do Sul, também tem uma forte base industrial, especialmente nas áreas de metalurgia, maquinário e agroindústria. Não obstante, o Sul também apresenta IDHM mais alto em comparação com o Norte e o Nordeste. Isso pode estar relacionado à presença de instituições de ensino superior e técnicas de qualidade nessas regiões, o que pode explicar sua presença dominante entre as sete (7) ICTs-MACE.

Adicionalmente, o Nordeste tem experimentado um crescimento no setor industrial nas últimas décadas, mas ainda está atrás do Sudeste e Sul em termos de desenvolvimento tecnológico e industrial.

Outro fator exógeno para a constatação de sete (7) ICTs-MACE aponta para os resultados alcançados pela Política de Expansão da RFEPCT (2005-2015). Dentre esses resultados até 2016 em relação à expansão da RFEPCT, torna-se evidente que houve um foco em fomentar o desenvolvimento socioeconômico das regiões Norte, Nordeste e Centro-Oeste. Isso se alinha com a diretriz de equilibrar o desenvolvimento educacional e tecnológico no país, considerando as disparidades regionais.

Essa evidência pode ser constatada indiretamente, ao se analisar os resultados desse estudo, comparando-se as ICTs em relação aos seus resultados vinculados aos canais de PD&S, uma década após sua implementação enquanto política (Brasil, 2008), com duas ICTs dessas regiões aparecendo no resultado geral apurado entre 2008 e 2018, conforme já mencionado acima. Essa tendência sugere um crescimento robusto ao longo do tempo e destaca o Nordeste como uma região emergente no SNI.

Apesar dos esforços concentrados do governo em impulsionar o desenvolvimento socioeconômico nas regiões Norte, Nordeste e Centro-Oeste, a presença dominante das regiões Sudeste e Sul entre as sete (7) ICTs-MACE indica que ainda há desafios a serem enfrentados e reflete a concentração de desenvolvimento tecnológico, industrial e educacional nessas regiões, bem como reflete sua dominância em termos de qualidade e desempenho, enquanto ICTs.

No entanto, é factível afirmar que os investimentos e ações significativas na RFEPCT começaram a apresentar seus primeiros resultados positivos, em um curto

intervalo de 10 anos de funcionamento neste modelo, quando comparado às universidades federais que possuem muito mais anos atuando também no eixo do desenvolvimento e fomento do SNI, através da CT&I e PD&I.

Isso é corroborado tanto pelos fatores identificados nos relatórios do MEC/SETEC, CGU, CNPq, CAPES, dentre outros (Analytics, 2018; 2019; Brasil, 2021c), quanto pelos resultados encontrados neste estudo, no qual foi possível constatar que a região Nordeste está presente entre as sete (7) ICTs-MACE, já na primeira década de atividades intensificadas da RFEPCT.

Os dados e resultados analisados evidenciam que, entre 2008 e 2018, as ICTs das regiões Nordeste e Centro-Oeste têm aumentado sua contribuição ao SNI. Esse crescimento destaca a eficácia da política de expansão da RFEPCT em democratizar o acesso à educação e fomentar o desenvolvimento regional. Essa ênfase na expansão regional é uma diferença notável em relação ao estudo de Perucchi *et al.* (2015), que não abordou especificamente a expansão geográfica das instituições de pesquisa e inovação.

As regiões Norte e Nordeste foram priorizadas devido à demanda reprimida por educação profissional e tecnológica, bem como pela necessidade de impulsionar o desenvolvimento econômico e social, prioridade essa refletida entre as sete (7) ICTs-MACE, contidas no “*cluster 1*” (HC e NHC, Abordagem 2). Contudo, é essencial enfrentar os desafios pendentes para assegurar que todas as regiões do Brasil disponham de ICTs de alta qualidade com desempenho equiparado.

A partir da análise minuciosa do “*cluster 1*” e das ponderações individuais de cada uma das seis (6) IPCT analisadas (IFES, IFCE, IFSP, CEFET/MG, IFRS e IFRN), foi possível delinear os padrões e relações existentes entre seus canais de PD&S e sua contribuição ao modelo TH de inovação, especialmente no que tange ao “*cluster 1*”, abordando os possíveis fatores endógenos pertinentes.

O IFSP desponta notavelmente no Componente Principal 1, que corresponde aos “Canais de Pesquisa e Desenvolvimento Tradicionais”, apresentando uma pontuação de 3,23, a mais alta entre as instituições avaliadas, evidenciando seu papel central na inovação e no avanço de pesquisas nesse segmento. Além disso, o IFRS se destaca nesse fator com uma pontuação de 2,20, demonstrando, assim como o IFSP, uma base sólida em pesquisa e desenvolvimento. Essas constatações, baseadas nos dados, reforçam a relevância contemporânea dos métodos tradicionais de pesquisa, sugerindo uma forte contribuição dessas instituições na dimensão universitária da inovação TH, na qual a geração de conhecimento científico e tecnológico é um fator relevante à inovação.

Adicionalmente, comparando o IFSC com o IFSP e o IFRS, foi possível constatar que ele possuía uma pontuação de 1,05, na posição mediana do IIP\_PD&S, o que demonstrou uma presença sólida, mas sem liderança destacada no desenvolvimento tradicional de pesquisa no resultado consolidado entre 2008 e 2018. Esses resultados são semelhantes ao estudo de Perucchi *et al.* (2015), no sentido de que ambos destacam a importância das atividades de pesquisa nas instituições educacionais. No entanto, o estudo atual vai além, fornecendo uma análise mais detalhada das diferentes instituições e seus desempenhos em vários Fatores.

O IFSC se destaca notadamente na cooperação com os SPP, evidenciado pela sua pontuação elevada de 5.88 no Componente Principal 2, que corresponde aos “Canais de Cooperação em Inovação com SPP”. Essa constatação evidencia uma colaboração e integração intensivas com esses serviços, consolidando o IFSC como líder incontestável neste domínio. Tal êxito sugere não apenas uma cooperação robusta, mas também uma abordagem possivelmente inovadora em suas estratégias de atuação junto aos SPP, contribuindo assim para uma interação mais dinâmica e produtiva no contexto do modelo TH de inovação.

O destaque do IFSC, como líder na cooperação com os SPP, contrapõe-se ao resultado obtido por Perucchi *et al.* (2015) que ressaltou a falta de motivação dos entrevistados para desenvolver pesquisas relacionadas às demandas do setor produtivo. Essa diferença indica que, ao longo dos anos, houve mudanças na abordagem das instituições, com algumas agora demonstrando maior envolvimento com o setor produtivo, o que é um sinal positivo de alinhamento com as necessidades do mercado.

O IFES e o IFRN destacam-se significativamente no Componente Principal 3, intitulado “Canais de Iniciativas de Propriedade Intelectual e Captação de Investimento”, registrando pontuações de 2,30 e 1,27, respectivamente, o que sinaliza um engajamento robusto em iniciativas ligadas à propriedade intelectual e à atração de investimentos. A análise constatou ainda que o CEFET/MG segue uma orientação similar, alcançando uma nota expressiva de 4,90 nesse mesmo fator, indicativo de uma postura proativa na busca por recursos e no desenvolvimento de projetos inovadores na esfera da propriedade intelectual. Esse perfil estratégico compartilhado por essas instituições não apenas as posiciona como referências em inovação e empreendedorismo dentro da RFEPCT, como pode facilitar a criação de parcerias e colaborações com entidades governamentais e industriais, potencializando o modelo TH, através da convergência de esforços e recursos.



Ao comparar o IFSC com o IFES e o IFRN, pode-se constatar que o primeiro não se destaca nesse fator, porém não deficitário, pois apresenta uma pontuação de -0,68. Por apresentar um excelente resultado no Componente Principal 2, reavaliar suas estratégias neste domínio pode potencializar seu desempenho, captando mais recurso, e consolidando-se como uma referência também neste fator.

O IFCE emerge com significativa predominância no Componente Principal 4, atrelado aos "Canais de Projetos Colaborativos Financiados pelos SPP", apresentando uma pontuação de 5,39 o que sinaliza uma robusta predisposição institucional para a cooperação e a realização de iniciativas conjuntas financiadas pelos SPP. Tal abordagem, dentro do escopo do modelo TH, não apenas fomenta a integração e a colaboração entre esferas governamentais, acadêmicas e industriais, mas também catalisa inovações colaborativas e a transferência de conhecimento, consolidando a posição de liderança do IFCE nesse quesito. Em contraste, o IFSC registra uma pontuação de -0,10 no mesmo fator, revelando uma margem considerável para aprimoramento em suas iniciativas de projetos colaborativos.

No que tange ao Componente Principal 5, associado aos "Canais de Consultoria e Treinamento para SPP", o IFSC encontra-se em um estágio inicial, marcando -0,19, um desempenho consideravelmente inferior, se comparado ao líder no domínio, que é o IFPE, seguido pelo IFPB, os quais alcançam uma pontuação expressiva de 3,74 e 2,73, respectivamente. Esses dados evidenciam a necessidade premente do IFSC em investir na expansão e fortalecimento de suas iniciativas de consultoria e treinamento, de modo a se aproximar dos padrões estabelecidos por instituições referenciais nesse fator. O IFPE e IFPB se firmam como uma referência, demonstrando uma trajetória consistente e promissora, o que reflete sua potente inclinação institucional para excelência nesse campo, ao passo que aponta para um horizonte de crescimento e aperfeiçoamento para o IFSC.

Por fim, em relação ao Componente Principal 6, que engloba aos "Canais de Atividades Diversificadas para SPP", o IFSC tem uma pontuação de 0,09. Esse resultado indica uma performance moderada. Constatou-se, no geral, que houve um equilíbrio entre as demais instituições nesse aspecto.

O desempenho do IFSC, ao ser comparado com as seis instituições que integram o "cluster 1" da RFEPCT, revela que sua maior força reside no Componente Principal 2. Neste aspecto, o IFSC se destaca, evidenciando uma colaboração intensa e frutífera com os SPP, algo que não só deve ser preservado, mas também potencializado.

Contudo, quando se analisa os Componentes Principais 3, 4 e 5, foi possível constatar que o IFSC apresenta um desempenho mediano ou até abaixo da média, comparativamente, entre as seis ICTs em destaque, a saber: IFES, IFCE, IFSP, CEFET/MG, IFRS e IFRN. Essa constatação sinaliza que há potencial para avanços significativos, mediante uma reavaliação estratégica cuidadosa, que objetive fortalecer a atuação e eficiência do IFSC nesses segmentos específicos.

Ao comparar o resultado dos *clusters* da Abordagem 2 de HC e NHC, de forma genérica, foi possível constatar que para além do “*cluster 1*”, comum a ambas, elas produziram agrupamentos distintos e com alta variabilidade. Neste contexto específico, um ARI de 0,27 sinaliza uma concordância limitada entre os agrupamentos gerados pelas análises HC e NHC, implicando que esses métodos estão discernindo padrões apenas parcialmente sobrepostos nos dados analisados. Ou seja, essa constatação sugere para uma alta variabilidade nos grupos identificados para além do “*cluster 1*”, que é comum em ambas as abordagens. Constatou-se também que a maioria das ICTs está agrupada sob os “*clusters 2 e 3*”, tanto em abordagens hierárquicas quanto não hierárquicas. Alguns casos apresentaram discrepâncias entre os agrupamentos HC e NHC, como IFSC e IFMA.

Em resumo, a análise das sete (7) ICTs-MACE evidencia uma pluralidade de focos e competências entre as instituições da RFEPCT, caracterizando uma valiosa diversidade, logo heterogeneidade. Esta última, por sua vez, confere à rede uma capacidade multifacetada de responder aos desafios que se apresentam no campo da educação profissional e científica no Brasil, destacando a importância de se valorizar e fomentar as distintas inclinações e forças de cada instituição dentro da rede.

A análise constatou uma significativa diversidade nas estratégias e níveis de maturidade das diferentes instituições envolvidas em uma variedade de iniciativas. Para um escrutínio mais detalhado, sugere-se a correlação dessas variáveis fatoriais com outros dados institucionais como financiamento e quantidade de estudantes.

Dentro dos agrupamentos hierárquicos e não hierárquicos, percebeu-se uma distribuição heterogênea das instituições, indicando que existem características específicas que influenciam essas classificações e que são capturadas pela análise fatorial através do PCA. Uma melhor compreensão dos critérios e parâmetros que embasam essa técnica de agrupamento poderia oferecer *insights* mais profundos e direcionar focos específicos de melhorias para cada instituição, visando um avanço no índice ora proposto IIP\_PD&S.

As ICTs apresentam uma diversidade geográfica notável, marcando presença abrangente em território nacional e consolidando polos de excelência em diversas regiões do Brasil. Nota-se, no entanto, um desempenho proeminente das instituições situadas nas regiões Sudeste e Sul, conforme indicado pelo índice IIP\_PD&S e corroborado pelas análises de agrupamento realizadas.

Para uma análise mais abrangente e uma perspectiva comparativa, foram consultados estudos semelhantes, como o de Arcuri (2016) que explorou as políticas de CT&I e financiamento público à infraestrutura de ciência e tecnológica (inter)nacionais, e De Negri, F. *et al.* (2016), que mapearam a infraestrutura científica e tecnológica brasileiras. Da mesma forma, o trabalho de Villasenor *et al.* (2017) trouxe *insights* sobre a inovação aberta nos SPP multinacionais, enquanto Miranda *et al.* (2018) investigaram a diversidade operacional entre as ICTs da RFEPCT. Berbegal-Mirabent *et al.* (2019) focaram na heterogeneidade das ICTs espanholas. Por fim, o estudo recente de Andrade *et al.* (2022) contribuiu com análises contemporâneas sobre a interação entre ICTs-GOV-SPP. Esses estudos, juntos, oferecem um panorama diversificado e complementar, considerando as especificidades e contextos variados de cada pesquisa.

O estudo de Miranda *et al.* (2018) agrupou as ICTs brasileiras em 2016, com base em suas características operacionais, classificando-as como Pequeno, Intermediário e Grande Porte. Esta categorização evidenciou que o tamanho de uma instituição não necessariamente determina sua eficiência. Em comparação, o presente estudo proposto, ao analisar a RFEPCT, constatou o IFRN, o IFCE e o IFSC como instituições de destaque, corroborando com a classificação de "Grande Porte" (Miranda *et al.*, 2018). Silva (2015) também constatou o bom desempenho do IFSC, IFGO, IFES, IFCE e IFRS, em relação às linhas de apoio de PD&I e extensão tecnológica. Isso sugere que essas instituições têm sido mais ativas na condução de pesquisas e atividades de inovação, indicando consistência nos resultados ao longo do tempo.

Villasenor *et al.* (2017) voltaram seu olhar para as ICTs mexicanas, concentrando-se na análise cienciométrica. Ambos os estudos, tanto o mexicano quanto o presente, valorizam a tomada de decisão baseada em dados e ressaltam a importância de ferramentas analíticas na avaliação do desempenho das ICTs.

A pesquisa de Andrade *et al.* (2022) adentrou nas capacidades inovadoras de 8 ICTs brasileiras, adotando ferramentas estatísticas como PCA e HC para esboçar o TH. Adicionalmente, a esta análise, o presente trabalho avançou na metodologia ao incorporar a NHC e a propor uma análise simultânea com as três técnicas (HC, NHC e EFA com

PCA) dos resultados, a partir da mineração dos dados obtidos. Some-se, a essa adição, a disponibilidade de uma ferramenta tecnológica *online* para monitoramento em tempo real dos resultados dos canais de PD&S das ICTs.

Ambos os estudos constataram uma taxa preocupante de transferência de tecnologia no Brasil. Enquanto Andrade *et al.* (2022) destacou os valores baixos obtidos de sua análise para regiões menos desenvolvidas, a análise da RFEPCT constatou taxas similarmente baixas de transferência de produções tecnológicas das IEPTC da RFEPCT para os SPP, independentemente da região geográfica. Somam-se a essas constatações as preocupações já compartilhadas por Perucchi *et al.* (2015) sobre a baixa taxa de transferência de tecnologia no Brasil. Isso sugere que, apesar dos esforços em pesquisa e desenvolvimento, há um desafio persistente em traduzi-los em inovações que beneficiem os SPP. Corroboram-se com os estudos sob esse tema, pois é necessário enfatizar a necessidade de melhorias nesse aspecto.

Perucchi *et al.* (2015) destacam a importância do financiamento público para as atividades de pesquisa e inovação nas instituições de ensino técnico e tecnológico, Andrade *et al.* (2022) também ressaltam a relevância do financiamento público nas atividades de C&T e inovação das ICTs, assim como De Negri, F. *et al.* (2016) e Arcuri (2016), e introduziram uma estrutura de análise que integra o modelo TH. Como extensão a análise de Andrade *et al.* (2022), esta tese foi um passo além, ao propor um modelo analítico integrado, que incorpora técnicas de análise multivariada para gerar *insights* sobre vários aspectos relacionais aos canais de PD&S e de interação com modelo TH de inovação. Ressalva-se, entretanto, que, este estudo constatou que mais de 99% da produção científica é financiada por recursos públicos, indicando uma dependência significativa do financiamento governamental. Isso é uma preocupação, pois a diversificação das fontes de financiamento é importante para garantir a sustentabilidade das atividades dos canais de PD&S.

Em relação às colaborações com os setores produtivos, Andrade *et al.* (2022) realçaram as associações com grandes empresas, como a Petrobras. Por outro lado, observa-se que a RFEPCT ainda possui parcerias limitadas com empresas desse porte. Quanto às variáveis relevantes, Andrade *et al.* (2022) identificam projetos de pesquisa de doutorado e patentes como primordiais, cobrindo 93% da variabilidade total, que ao analisarem os resultados dos canais de PD&S da RFEPCT orientações de pesquisa ( $X_3$ ), projetos de pesquisas em cooperação com SPP ( $X_8$ ) e patentes publicadas ( $X_{20}$ ), abrangem 40,2% da variabilidade total.

Em resumo, Andrade *et al.* (2022) ofereceu uma visão detalhada das capacidades inovadoras de 8 ICTs brasileiras. Em contraste, esta tese abrange a população de 42 ICTs da RFEPCCT. Além disso, ela introduz uma ferramenta tecnológica *online* baseada em BI. Essa solução permite que o usuário customize sua análise, conforme o grau de especificidade dos dados que busca. Tal ferramenta atende uma necessidade dos gestores das ICTs: dispor de um recurso que responda prontamente às demandas dos órgãos de controle brasileiros, fornecendo relatórios e informações auditadas e validadas institucionalmente.

No que diz respeito à infraestrutura e financiamento, De Negri, F. *et al.* (2016) e Arcuri (2016) analisaram o sistema brasileiro de C&T, enfatizando a necessidade de melhorias na qualidade, relevância internacional e diversificação das fontes de financiamento. Eles também ressaltaram o papel central do Estado no fomento à pesquisa e inovação. Ao analisar os resultados da RFEPCCT, neste contexto, fica evidente que essa realidade ainda era presente em 2018, pois mais de 99% da sua produção científica, sob a forma de orientação de pesquisa, por exemplo, foram financiadas por recursos públicos, realidade essa que ainda deve ocorrer atualmente (consultar MVP, Módulo Amplo, Apêndice B, Figura B. 15).

Berbegal-Mirabent *et al.* (2019) analisaram a heterogeneidade das universidades na Espanha, destacando a especialização institucional como um determinante-chave de desempenho. A pesquisa identificou caminhos distintos de desempenho entre universidades e a importância do ambiente geográfico e regulatório. Esses estudiosos, considerando o contexto e analisando os resultados da RFEPCCT, suas ICTs, entre 2008 e 2018, também seguiram caminhos distintos de desempenho, quando comparadas ao índice aqui proposto, corroborando com as constatações de heterogeneidade, ambiente geográfico e regulatório, constatado no caso das ICTs espanholas.

O estudo atual reconhece a heterogeneidade das instituições de ensino profissional e tecnológico no Brasil, destacando a importância de valorização e de fomento às diferentes inclinações e forças de cada instituição, dentro da rede. Isso se alinha com a constatação de Perucchi *et al.* (2015) de que o tamanho de uma instituição não, necessariamente, determina sua eficiência.

O estudo atual fornece uma análise mais abrangente e detalhada das instituições de ensino técnico e tecnológico no Brasil, em relação à pesquisa, desenvolvimento e inovação. Ele destaca mudanças na abordagem das instituições ao longo dos anos e enfatiza a importância da diversificação de fontes de financiamento, cooperação com o

setor produtivo e transferência de tecnologia. Essas descobertas são consistentes com as preocupações levantadas no estudo de Perucchi *et al.* (2015) e destacam a necessidade contínua de melhorias nas políticas e práticas de pesquisa e inovação no país.

O estudo atual destaca que as instituições não se limitam à produção de artigos científicos, mas também que se envolvem em outros canais de PD&S. Perucchi *et al.* (2017) observaram que apenas 17% dos docentes analisados em seu estudo tinham produção em artigos, sugerindo que as instituições devem incentivar essas atividades.

O texto menciona a importância das parcerias e colaborações interinstitucionais no contexto da TH. Isso está alinhado com as conclusões de outros estudos que enfatizaram a colaboração como um fator-chave para o sucesso em pesquisa e inovação. Essa ênfase na colaboração é crucial para fortalecer a interação entre academia, setor produtivo e governo.

Miano *et al.* (2020) observaram a necessidade de indicadores mais precisos para avaliar a produção científica dos grupos de pesquisa nas instituições. Essa é uma questão importante, pois indicadores precisos são essenciais para acompanhar as políticas e as ações implementadas para promover o fomento dos canais de PD&S e a solução tecnológica aqui proposta busca contribuir para mitigar essa necessidade.

No que se refere a Rossi (2022), o estudo atual concorda com a constatação de que a falta de recursos humanos especializados na transferência de tecnologia pode ser um obstáculo à inovação. Ambos os estudos destacam a importância de recursos humanos capacitados para facilitar a transferência de tecnologia e promover a inovação. Isso enfatiza a necessidade de investimento em capacitação e treinamento nessa área.

Quanto aos estudos de Miano; Couto; *et al.* (2017), Miano *et al.* (2020) e Ribeiro *et al.* (2020), o estudo atual não abordou diretamente a atipicidade dos grupos de pesquisa ou a trajetória quantitativa e qualitativa da produção científica deles. Portanto, não há uma comparação direta com esses estudos em relação aos resultados específicos.

No geral, o estudo atual se baseia em pesquisas anteriores, como Silva (2015), Perucchi *et al.* (2017), Miranda *et al.* (2018), Rossi (2022), Miano; Couto; *et al.* (2017), Miano *et al.* (2020) e Ribeiro *et al.* (2020), para oferecer uma análise mais abrangente e atualizada do estado das ICTs da RFEPCT, em relação aos seus canais de PD&S. Assim, como eles, destacam-se os desafios persistentes, como a baixa produção técnica e de inovação, a falta de registros de patentes e a importância de recursos humanos especializados na transferência de tecnologia. Essas conclusões são consistentes com os

estudos anteriores e destacam áreas de melhoria contínua no sistema de educação técnica e tecnológica do país.

Em síntese, as interações TH são cruciais para a inovação, e as estratégias internas das ICTs determinam seus caminhos de desempenho. Tanto os contextos nacionais quanto os internacionais influenciam a performance e as capacidades inovadoras das ICTs; a combinação de políticas eficazes, investimentos estratégicos e colaborações interinstitucionais pode potencializar o impacto dessas instituições na inovação global.

### **5.7 Análise dos Canais de Pesquisa, Desenvolvimento Tecnológico e Serviços em Ecossistema da Rede de Formação e Pesquisa em Ciência e Tecnologia: um Estudo de Caso Aplicado à Plataforma Proposta**

Utilizando dados coletados de diversas fontes, a plataforma, neste estudo proposto, visa embasar o processo de tomada de decisão, análise, gerenciamento e acompanhamento das demandas das ICTs, SPP e GOV. Funcionando como um Sistema de Informação Gerencial, de Suporte à Decisão e de Suporte Executivo, a solução foi projetada para se adaptar ao nível de acesso e configuração visual de cada usuário.

Diferentemente de outras soluções existentes, o MVP proposto preenche lacunas, agregando múltiplas fontes e múltiplos dados relacionadas aos canais PD&S em uma plataforma única. Foi projetada para permitir o acesso aberto, público, transparente e auditável aos resultados desses canais. Adicionalmente, disponibiliza uma interface intuitiva e permite uma variedade de análises e filtros, possibilitando um vasto leque de comparações entre as variáveis apresentadas ao longo do tempo, conforme a disponibilidade e detalhamento/granularidade dos dados acessíveis e passíveis de validação.

Conforme mencionado acima, o MVP proposto propicia análises em diferentes níveis de detalhamento, como: Amplo, Detalhado e Comparativo. Isso ilustra a variação no número resultados e soluções disponíveis pelos canais de PD&S das ICTs. Isso evidencia as inúmeras possibilidades de interação com os setores produtivo, acadêmico e público. Por exemplo, a análise dos dados levanta questionamentos sobre o tipo de consultoria que realizada e cadastrada e, como o MVP proposto permite explorar essa informação.

Dentre os recursos disponíveis, por exemplo, o filtro de texto permite ao usuário acessar informações detalhadas sobre consultorias específicas, por meio dos mapas de nuvens de palavras, evidenciando a utilidade da plataforma para uma análise mais

profunda dos dados, coletando informações preliminares para planejar um estudo de *text mining*<sup>25</sup> (Tan, 1999) mais detalhado em outra ocasião. No entanto, seria benéfico discutir de forma mais abrangente as implicações dessas descobertas e como elas se relacionam com pesquisas anteriores, como o estudo de Andrade *et al.* (2022).

A plataforma proposta permite aos usuários identificar, por exemplo, que tipo de orientação de pesquisa foi publicada, a qual área ela foi desenvolvida e quais regiões do Brasil pesquisas similares foram desenvolvidas (Subseções 4.3 e 4.4). Ela também oferece a opção de cruzar dados, como a titulação máxima dos pesquisadores, para entender se a titulação, como outro exemplo, favorece a oferta de consultorias, projetos e orientações de pesquisa concluídos, em andamento ou desativados. Essas e outras comparações podem ser utilizadas para realizar análises mais complexas a depender do interesse do usuário (Subseção 4.4).

O MVP, aqui proposto, possui 3 módulos, com algumas funcionalidades semelhantes entre si, tais como: a disponibilidade de gráficos de vários tipos; porém, é altamente especializada e restrita aos canais de PD&S, enquanto a plataforma proposta por Azevedo *et al.* (2021) apresenta uma série de 4 módulos, com indicadores, filtros e gráficos, para cada um deles, direcionados ao ensino, pesquisa, extensão e gestão em uma única ICTs.

Diferentemente das plataformas citadas, a proposta desta tese permite comparações espaço-temporais entre instituições, unidades federativas, regiões e até países, e é alimentada por múltiplas fontes de dados públicas. Azevedo *et al.* (2021) não definiram explicitamente a origem dos dados coletados, o ano de coleta etc., enquanto os dados da NTDATA são fornecidos, principalmente, por formulários encaminhados aos diversos NIT do nordeste. Além disso, em relação aos dados geográficos, o presente trabalho permite a aplicação de resultados levando-se em consideração mapas estaduais, regionais ou nacionais, demonstrando a concentração, ou não, de determinadas ações e trazendo uma visão semelhante à apresentada por Johnson *et al.* (2022).

---

<sup>25</sup> O *text mining*, também conhecido como mineração de texto, é um processo que envolve a extração de informações relevantes e significativas de grandes volumes de dados textuais. Utilizando técnicas de processamento de linguagem natural (NLP), análise estatística e aprendizado de máquina, o *text mining* busca identificar padrões, tendências e relações ocultas nos dados, transformando-os em informações úteis e estruturadas para tomada de decisões e análise. Essa abordagem é amplamente utilizada em diversos campos, como análise de sentimentos, pesquisa científica, inteligência de mercado e monitoramento de redes sociais.



O módulo **Detalhado**, após selecionar a métrica “*Orientação de Pesquisa*”, como exemplo, disponibiliza outros resultados organizados e com um maior detalhamento para essa métrica. Nesse módulo, por exemplo, considerando o período de 2008 a 2018, seis ICTs foram responsáveis por 28,46% das orientações de pesquisa, entre 41 instituições da RFEPCT analisadas. Essas orientações foram conduzidas principalmente por mestres (56%), seguidos por especialistas (27%) e doutores (12%), num total de 55.256 pesquisadores. Entre os orientadores, 92% foram dedicados exclusivamente por contrato. Apenas 23,84% dessas orientações receberam financiamento de bolsas e, entre aquelas que receberam, apenas 1,3% foram classificadas como relevantes por seus guardiões, 59,8% foram financiadas pelo SPP e 39,8% pelo ICTs e GOV, sendo esse último financiamentos relacionados aos recursos públicos.

A análise do MVP proposto também permitiu constatar que, de 2015 a 2018, a quantidade de “orientação de pesquisa” ( $X_3$ ) caiu drasticamente, possivelmente devido à supressão dos recursos públicos ( $X_1$ ) nesse período. Embora não seja o objetivo deste trabalho, essa hipótese pode ser confirmada através de um estudo correlacionando variáveis disponíveis no módulo PD&S. Esse exemplo evidencia o grande número de informações e combinações possíveis a serem extraídas, levando-se em conta ano, região, unidade federativa, instituições pesquisadas e as 24 dimensões e suas subdivisões.

Outro exemplo de dados que podem ser agregados está relacionado aos setores da economia brasileira, conforme regulamentado pela Classificação Nacional e Atividades Econômicas do Brasil (CNAE), a Nomenclatura de Atividades Econômicas da União Européia (NACE) ou o Sistema de Classificação da Indústria da América do Norte dos EUA (NAICS). Essas classificações podem ajudar a identificar quais setores podem ser beneficiados e qual a tendência, por exemplo, das soluções de pesquisas desenvolvidas nas ICTs. Com isso, em novo exemplo, é possível comparar se as pesquisas desenvolvidas estão proporcionalmente distribuídas à quantidade de CNAE (Quadro B. 16, Apêndice B) da região ou se os projetos de PD&S estão alinhados aos planos de desenvolvimento socioeconômico. O Observatório de Ciência, Tecnologia e Inovação criou KPIs relacionados aos serviços mais intensivos em conhecimento, compreendendo as seções J, K, M, Q e R da CNAE (Brasil, 2020), também usadas em estudos que investigam a sinergia da TH (Leydesdorff *et al.*, 2019; Porto-Gomez *et al.*, 2019).

A análise conceitual mencionada acima considera os KPIs relacionados aos serviços mais intensivos em conhecimento, abrangendo as seções J, K, M, Q e R,

comparando essa relação com os resultados da RFEPCT, sendo possível observar algumas correlações e diferenças.

Por exemplo, na RFEPCT, as seções J (Informação e Comunicação), M (Atividades Profissionais, Científicas e Técnicas) e Q (Saúde Humana e Serviços Sociais) apresentam valores de representação, com (3,7%), (16,4%) e (4,08%), respectivamente. Estes valores sugerem que as *Orientações de Pesquisas* estão alinhadas com a natureza intensiva em conhecimento. A seção K (Atividades Financeiras, de Seguros e Serviços Relacionados) possui um valor de (0,27%), enquanto a seção R (Artes, Cultura, Esporte e Recreação) apresenta um valor de (0,22%).

Por outro lado, a RFEPCT destaca-se na seção P (Educação), com um valor de representação consideravelmente alto, de 38,2%. Essa área não é diretamente mencionada na análise conceitual dos KPIs, mas é fundamental para a formação de recursos humanos capacitados e para o desenvolvimento de pesquisas e inovações, que são aspectos centrais da sinergia da TH.

Em resumo, os resultados da RFEPCT, disponíveis no módulo **Amplio** do MVP proposto apresentam uma sinergia com as áreas intensivas em conhecimento mencionadas na análise conceitual, mostrando destaque nas seções J, M e Q. A seção P (Educação) também é fundamental para o contexto da RFEPCT, embora não esteja incluída na análise dos KPIs. Essa comparação pode ajudar a identificar áreas de atuação e colaboração, bem como oportunidades de desenvolvimento e crescimento para a RFEPCT em ecossistema de inovação e conhecimento.

Durante os testes funcionais (Subseção 4.3) o MVP proposto mostrou-se uma solução relevante e inovadora para apoiar os gestores como um sistema de informação gerencial, de apoio à decisão e/ou executivo, em consonância com os princípios básicos mencionados por Stair *et al.* (2010), de ter uma interface gráfica amigável, embora alguns especialistas tenham feito sugestões de melhorias para versões futuras.

Os resultados obtidos, a partir da organização de dados e atributos métricos em um painel baseado na arquitetura de BI proposta, permitem que gestores e pesquisadores sugiram conjuntos de regras aderentes às suas necessidades. Essas regras podem ser desenvolvidas a partir de novas pesquisas e, uma vez validadas, adicionadas à estrutura de BI proposta.

Para este estudo, foram trabalhados dados de várias fontes públicas, permitindo a visualização deles, referentes aos canais de PD&S. De forma similar, estudos como os de Jovanović *et al.* (2022), Azevedo *et al.* (2021), Berbegal-Mirabent *et al.* (2019),

Fernandez *et al.* (2017), Khatibi *et al.* (2017), Di Tria *et al.* (2015), Zhai *et al.* (2010), Guitart *et al.* (2015), dentre outros, também investigaram pelo menos uma das métricas estudadas neste trabalho no contexto das ICTs.

As questões mais importantes decorrentes da arquitetura proposta podem ser resumidas da seguinte forma: i) este MVP pode ser aplicado a qualquer ICT e foi apresentado por meio de um painel que pode ser expandido e transformado em aplicativo móvel e/ou *online*, como já ocorreu nas plataformas Nilo Peçanha, Universidade 360, dentre outras. A solução aqui proposta foi planejada e projetada para atender usuários interessados em monitorar, avaliar e encontrar soluções de oportunidade em ecossistema das ICTs. O MVP proposto utiliza informações públicas, sem conflito com a lei nº 12.527/2011 (Brasil, 2011) sobre confidencialidade de dados públicos.

A plataforma, ora proposta, não se encerra em si mesma, identificando lacunas de dados que podem ser remediadas para torná-los mais confiáveis e auditáveis. Algumas medidas incluem: criar atributos objetivos relacionados à captação de dados; auditar e validar informações nos currículos Lattes dos servidores vinculados às ICTs, principalmente, às federais; facilitar o acesso às informações sobre investimentos públicos e/ou privados para os canais de PD&S, categorizados e registrados; descrever metadados referentes aos atributos textuais utilizados durante o registro das atividades para os canais de PD&S.

Tal proposta visa não apenas organizar e apresentar os dados relacionados às atividades e aos dados de PD&S em ecossistema das ICTs, mas também identificar e abordar lacunas existentes nas soluções atuais. Por meio da análise de dados e da implementação de regras e métricas mais aderentes às necessidades dos gestores e pesquisadores, este painel pode contribuir significativamente para a tomada de decisões e para a proposição de políticas e ações voltadas para a inovação e a sinergia da TH. O modelo de BI sugerido possibilita a ampliação e adaptação do MVP proposto a variados contextos e instituições, valorizando-a como instrumento para monitorar, avaliar e compreender produtos, processos e serviços de PD&S e seus efeitos no desenvolvimento socioeconômico (Subseção 4.3).

Dessa forma, o MVP proposto neste trabalho busca não apenas oferecer uma solução abrangente para gerenciar e apresentar informações relevantes, mas também funcionar como uma plataforma para melhorar continuamente a qualidade e a confiabilidade dos dados disponíveis. Ao abordar as lacunas existentes e adaptar-se às necessidades específicas de cada instituição e contexto, o MVP proposto pode

desempenhar um papel crucial na promoção da colaboração entre os atores da TH e no fomento de uma cultura de inovação e desenvolvimento sustentável.

Em síntese, ele consiste numa proposta nacional, inovadora e relevante, que atende às demandas das ICTs, no que diz respeito à análise e à compreensão dos aspectos cruciais de seus canais de PD&S (Apêndice B, Quadro B. 17). Embora tal solução necessite de melhorias, até porque é um MVP, com mais testes de mercado e seu lançamento para a comunidade avaliar por mais tempo, garantiria condições de ser aprimorado, a partir de discussões mais amplas, em relação as que foram avaliadas apontadas nos resultados. Ainda assim, o MVP proposto apresenta um avanço significativo em relação às soluções anteriores, sem similaridade até a presente data (Subseção 4.3 e 4.4 e Apêndice B, Quadro B. 3: Especialistas 1, 3, 4, 5 e 6).

### **5.7.1 Estudo de Caso Aplicado: Análise do Cenário Nacional de CT&I e PD&I na Primeira Década de Existência da RFEPCT**

O modelo clássico de BI pode apresentar vários tipos de resultados (Fana *et al.*, 2021; Krmac, 2011). O modelo empregado para este trabalho buscou exibir dados por meio de *dashboard* e usar a *Data Mining* para identificar padrões e *insights*, com base em dados que auxiliem a tomada de decisões (Subseção 3.2, 4.3 e 4.4, Apêndice B, Tabela B. 8). A seguir são apresentadas as análises e discussões resultantes do mapeamento dos dados dos canais de PD&S da RFEPCT, à luz do modelo clássico de TH (Cai *et al.*, 2021; Etzkowitz, 1996; Etzkowitz *et al.*, 2000), bem como de alguns indicadores públicos nacionais (Brasil, 2020; 2021a) e internacionais (Bank's, 2020) sobre qual as características do ecossistema de CI&T e PD&I, em que a RFEPCT se desenvolveu em sua primeira década de existência, nos moldes da lei de 11.892, 31 de dezembro de 2008 (Brasil, 2008).

Entre 2008 e 2018, o Brasil adotou uma política contrária à taxa de crescimento média *Year Over Year* (YOY)<sup>26</sup>, comparado à média do índice *Research And Development Expenditure* (% GERD) apresentado pelos BRICS (para efeito de cálculo comparativo, o Brasil foi excluído) e as cinco primeiras nações (em ordem crescente Suíça, Suécia, Reino Unido, Holanda, Estados Unidos – TOP 5) mais inovadoras,

---

<sup>26</sup> *Year Over Year* (YOY) – cálculo desenvolvido para comparar resultados em um determinado período com o mesmo período anterior, seu resultado é expresso em percentagem.

segundo o *Global Innovation Index*<sup>27</sup> (GII), entre 2013 e 2018, disponibilizado pelo banco mundial<sup>28</sup>.

Ao analisar o YOY, o dispêndio TOTAL em P&D<sup>29</sup>, em relação ao percentual do Produto Interno Bruto (PIB) – P&D (%PIB), foi possível concluir que os BRICS e os TOP 5 aumentam em média de 221,5% e 50%, respectivamente, em comparação ao Brasil. Em relação à participação do GOV no TOTAL, constata-se que o governo do Brasil aumentou sua participação em 270%, se comparado aos TOP 5 e 29% aos BRICS. Por fim, ao se analisar uma das metas das diversas políticas implementadas no período que era o aumento do dispêndio em P&D (%PIB), por parte dos setores produtivos (em inglês, *Business Enterprise Sector* - BES), constata-se mais uma vez que as ações não surtiram o efeito esperado. Ao ser comparada a diferença entre a participação do BES do Brasil com os TOP 5, constata-se que, enquanto o BES dos TOP 5 aumentou em 800% e os BES dos BRICS, 5.360% sua participação no total do dispêndio de P&D (%PIB), o Brasil, praticamente, manteve-se estável.

Embora o Brasil ocupe mundialmente o décimo terceiro lugar em publicações científicas, seu desempenho em outros indicadores mundiais vinculados à P&D sugerem uma possível dificuldade em transformar esse conhecimento em uma *commodity* com alto valor agregado.

Por exemplo, ao serem analisadas as bases de dados estatísticos do comércio exterior brasileiro de bens<sup>30</sup>, disponibilizado pelo Ministério da Indústria, Comércio Exterior e Serviços - MDIC, entre 2008 e 2018, pode-se constatar que os produtos brasileiros mais exportados são essencialmente oriundos do setor primário. Embora o país seja destaque mundial no desenvolvimento de P&D, principalmente no setor do agronegócio e minério, os *royalties* desses setores comparados aos dos produtos com alto valor tecnológico agregado são relativamente menores.

---

<sup>27</sup> O *Global Innovation Index* (GII) é um índice de referência mundial utilizado pelas organizações como ONU, Banco Mundial, OCDE, WIPO, dentre outros, quando o objetivo é mensurar o desempenho de inovação de uma economia, disponível no sítio eletrônico <https://www.globalinnovationindex.org/analysis-indicator>, acessado em março de 2019. Em 2008, os *top* dez (10) eram representados pelos seguintes países em ordem crescente: 1 Suíça, 2 Holanda, 3 Suécia, 4 Reino Unido, 5 Cingapura, 6 Estados Unidos da América, 7 Finlândia, 8 Dinamarca, 9 Alemanha, 10 Irlanda.

<sup>28</sup> A análise foi realizada a partir dos dados fornecido pelo banco mundial, disponível em <https://tcddata360.worldbank.org/topics>, acessado em março de 2019.

<sup>29</sup> A análise foi realizada a partir dos dados fornecido pelo banco mundial, disponível em <https://data.worldbank.org/indicator/GB.XPD.RSDV.GD.ZS?end=2018&locations=BR-KR-DE-FI-CH-IL-SG-SE-US-JP-FR&start=2008&view=chart>, acessado em março de 2019.

<sup>30</sup> Ministério da Indústria, Comércio Exterior e Serviços – MDIC, <https://dados.gov.br/dataset/bases-de-dados-estatisticos-do-comercio-exterior-brasileiro-de-bens>.

Esses resultados sugerem que as políticas adotadas pelo Brasil não foram eficazes em promover o crescimento em P&D, no período considerado, especialmente em comparação aos países BRICS e TOP 5. Isso destaca a necessidade de reavaliar e ajustar as políticas e estratégias brasileiras para incentivar a inovação e o desenvolvimento sustentável, tanto no setor público quanto no privado. Além disso, a análise reforça a importância do modelo clássico de Hélice Tríplice (TH) e do uso de indicadores públicos nacionais e internacionais para avaliar e orientar as políticas de P&D e inovação.

Um possível impacto desses resultados na evolução dos canais de PD&S da RFEPCT entre 2008 e 2018 pode ser observado em diversas áreas. A ineficácia das políticas adotadas pelo Brasil na promoção de crescimento em P&D sugere que houve pouca evolução positiva nos canais de PD&S durante esse período. A comparação desfavorável com os países BRICS e TOP 5 indica a necessidade de reavaliar as políticas e estratégias brasileiras para melhorar a competitividade e o desenvolvimento sustentável.

Outra constatação dos impactos das políticas pensadas para a promoção da P&D, no Brasil, pode ser examinada, ao se analisar a evolução do *ranking* brasileiro, entre 2008-2018, em onze (11) indicadores<sup>31</sup> mundiais, aderentes ao escopo desta tese, disponibilizados pelo Banco Mundial, os quais são utilizados para os planejamentos das políticas e ações das nações para o setor.

Considerando que a economia brasileira possui uma base primária, observou-se um retrocesso do Brasil no *ranking* internacional de inovação no período de 2008 a 2018. No critério *Quality of Scientific Research Institutions*, o país caiu 37 posições, passando da 43ª para a 77ª colocação em um conjunto de 151 países avaliados (N=151; da 43ª para 77ª); *High-Tech and Medium High-Tech Output*, caiu 8 posições (N=119; da 22ª para 30ª posição); *Capacity for Innovation*, caiu 46 posições (N=151; de 27ª para 73ª posição); *University/industry research collaboration*, caiu 25 posições (N=213; de 42ª para 67ª posição); e *University-Industry Collaboration in Research & Development*, caiu 20 posições (N=151; de 50ª para 70ª posição). Estes dois últimos evidenciam e corroboram os resultados constatados na participação dos setores produtivos no dispêndio nacional em P&D (% PIB).

---

<sup>31</sup> Banco Mundial, indicadores considerados: *Quality of Scientific Research Institutions*, *Human capital and research*, *Gross Expenditure On R&D (GERD)*, *High-Tech Exports*, *Global Innovation Index (I)*, *High-Tech and Medium High-Tech Output*, *Research and Development (R&D)*, *University/industry research collaboration*, *Capacity For Innovation*, *University-Industry Collaboration In Research & Development*, *Scientific and technical publications*.

Em relação à *Quality of Scientific Research Institutions*, deve-se considerar que em média 80% das instituições de ensino superior no Brasil são particulares e a grande maioria não é elencada nos *rankings* nacionais ou internacionais, o que possivelmente pode estar influenciando no resultado do índice supracitado. Todavia, também foram constatados, avanços do Brasil em ecossistema da inovação. No critério *Human capital and research*, subiu 23 posições (N=213; de 75<sup>a</sup> para 52<sup>a</sup> posição); *Gross Expenditure on R&D (%GERD)*, subiu 4 posições (N=142; de 31 para 27<sup>a</sup> posição); *High-Tech Exports*, subiu 9 posições (N=159; de 44<sup>a</sup> para 35<sup>a</sup> posição); *Research and Development (R&D)*, subiu 5 posições (N=211; de 33<sup>a</sup> para 28<sup>a</sup> posição); e *Scientific and Technical Publications*, subiu 3 posições (N=125; de 57<sup>a</sup> para 54<sup>a</sup> posição).

De acordo com o estudo apresentado (Barragan-Ocana *et al.*, 2020), os países mais inovadores são aqueles com maior investimento em pesquisa e desenvolvimento, maior número de patentes e maior número de publicações científicas. Portanto, é importante que o Brasil continue a investir em P&D e a promover políticas que incentivem a inovação no setor privado, a fim de melhorar sua posição no *ranking* global de inovação.

Considerando que 99% da P&D do Brasil ocorre nas ICTs públicas e que esse conjunto de índices estão mais direcionados às atividades fins dessas instituições, há indícios de que as políticas de expansão dessas instituições, estratégicas para o desenvolvimento nacional, de algum modo, podem ter influenciado no resultado ainda pouco relevante, por exemplo, em relação a 0,11%, do total de orientações de pesquisas (concluídas) em interação com os SPP por pesquisadores da RFEPCT, de 2008 a 2018. Portanto, pode-se inferir, com base nos resultados, que os esforços nacionais para se obter melhores resultados, ainda precisam ser muito fomentados, considerando o modelo econômico atual, baseado no conhecimento.

Os resultados, também, corroboram com a necessidade de sistematizar e monitorar as atividades, as condições e as diretrizes vinculadas ao desenvolvimento inovativo, baseados em dados, principalmente das ICTs públicas, as quais são o principal vetor estratégico da promoção da inovação e objeto desta investigação.

Embora seja constatado um *déficit* no investimento para o fomento da P&D e ações insuficientes de interação ICTs-SPP, a falta de progresso nos canais de PD&S pode gerar consequências de longo prazo, como redução do potencial de inovação, menor capacidade de competir no mercado global e menor crescimento econômico. A importância do modelo clássico de TH e do uso de indicadores públicos nacionais e internacionais para avaliar e orientar políticas de PD&I se torna mais evidente diante

desse cenário e, neste contexto, apresenta-se o IIP\_PD&S. Portanto, é necessário repensar e ajustar as políticas e estratégias brasileiras, incentivando a inovação, o desenvolvimento sustentável, e fortalecendo os canais de PD&S e a competitividade do país no cenário global.

Para potencializar as interações ICTs, SPP e GOV com os canais de PD&S com foco na maximização dos *royalties*, é necessário conhecimento prévio e sistematizado, subsidiado pelo monitoramento constante do ecossistema das ICTs, em relação aos investimentos públicos e privados destinados ao fomento da pesquisa e desenvolvimento. Dentre os modelos disponíveis para orientar os gestores na tomada de decisões, há a teoria da TH, que é mais aderente aos princípios e objetivos estabelecidos pela lei nº 11.892/08 (Brasil, 2008), que institui a RFEPCT e cria os Institutos Federais de Educação, Ciência e Tecnologia.

Conforme Seção 2 e Subseção 2.4, a teoria da TH é uma abordagem que enfoca a interação entre ICTs, SPP e GOV para promover a inovação e o desenvolvimento socioeconômico. Ela propõe que a colaboração entre esses três atores pode impulsionar a inovação de forma mais efetiva do que se cada uma atuasse isoladamente.

Quando se trata de análise de dados relacionados à inovação, a teoria da TH ofereceu o suporte teórico e analítico para entender o fenômeno da interação por ela proposta. Por exemplo, neste trabalho, utilizou-se a teoria da TH para analisar a relação entre a quantidade de patentes públicas pela RFEPCT de  $\cong 0,2\%$  das interações com os SPP, e que  $\cong 0,1\%$  foi o número orientações de pesquisas concluídas que tiveram parceria com os SPP, e que  $\cong 0,004\%$  foi a quantidade de consultorias contratadas pelos SPP, e de  $\cong 0,03\%$  pelo GOV para a promoção da inovação, demandada à RFEPCT desenvolvida por seus pesquisadores.

Essa abordagem ajuda, por exemplo a identificar oportunidades para fortalecer a colaboração entre as três entidades, como a criação de parcerias público-privadas ou a oferta de incentivos para empresas que trabalham com universidades em projetos de pesquisa e desenvolvimento. Além disso, a análise da TH pode ajudar a avaliar o impacto das políticas governamentais e/ou institucionais sobre as diretrizes e ações relacionadas ao fomento dos canais de PD&S, examinando como eles afetam a colaboração entre as entidades.

Em resumo, a teoria da TH foi útil na análise de dados relacionados aos canais de PD&S, conseqüentemente, à inovação, fornecendo o aporte teórico e analítico para



entender como a interação entre ICTs, SPP e GOV pode impulsionar a inovação e promover o desenvolvimento socioeconômico.

O MVP proposto oferece uma alternativa mais aderente às necessidades das ICTs para analisar, avaliar, interpretar e compreender os principais aspectos dos canais de PD&S (Subsecção 4.4.4). A metodologia utilizada para o desenvolvimento da solução é adequada (Subsecção 4.4.3) e já é comumente aplicada a dados educacionais e administrativos, como as plataformas Nilo Peçanha e Universidade 360 (Brasil, 2021a; 2022b).

A solução de BI adotada pode ser aplicada a sistemas de informação gerencial, de apoio à decisão ou à execução, e funcionalidades adicionais, bem como a uma API automatizada para carregar dados pós 2018; podem ser implementadas em versões futuras, conforme sugerido pelos especialistas.

A interseção entre as soluções de *dashboard* existentes e a solução apresentada é que ambas buscam organizar dados de ICTs ou disponibilizar indicadores para que ecossistemas, instituições ou mesmo nações possam desenvolver seus planejamentos utilizando fontes de dados públicas. No entanto, enquanto outras soluções concentram-se no monitoramento dos canais de ensino, aprendizagem, gestão, finanças e algumas específicas de pesquisa e/ou desenvolvimento de tecnologia, esta solução abrange os canais de PD&S integrados.

A depender do nível de acesso de cada usuário, o MVP pode ser integrada a sistemas de informação gerencial, de apoio à decisão ou à execução. Embora ela não implemente uma Interface de Programação de Aplicativos (em inglês, API) automatizada para carregar dados pós 2018 das fontes de dados usadas, tal funcionalidade pode ser implementada em versões futuras.

O Quadro B. 17 (Apêndice B) apresenta um resumo comparativo das funcionalidades e disponibilidade de informações entre MVP proposto e duas outras soluções. As principais funcionalidades foram comparadas com o objetivo de identificar as possíveis convergências e diferenças entre as propostas. Ao analisar criticamente as funcionalidades das quatro plataformas (o MVP proposto, Portal Integra, Portal Universidade 360 e *StelaExperta*<sup>®</sup>), pode-se observar que o MVP proposto oferece o maior número de funcionalidades, cobrindo a maior parte das necessidades dos usuários. A plataforma se destaca, principalmente, em aspectos como acompanhamento e monitoramento de performance, análise de métricas simultâneas, análise por Setor TH e visualização interativa de gráficos, mapas e tabelas.

O Portal Integra possui um número limitado de funcionalidades, em comparação com as outras plataformas, mas se destaca por oferecer cessão de infraestrutura, desenho industrial e desenvolvimento tecnológico. Além disso, é a única plataforma que oferece vitrine tecnológica institucional, bem como portfólio institucional.

O Portal Universidade 360, por sua vez, tem menos funcionalidades em comparação com o MVP proposto e *StelaExperta*<sup>®</sup>, mas oferece funcionalidades diferenciadas, como análise baseada em custos, tempo e qualidade, atualizações e adaptações contínuas e visão geral das métricas e KPIs.

A plataforma *StelaExperta*<sup>®</sup> possui um conjunto de funcionalidades bastante abrangente, semelhante ao MVP proposto, destacando-se principalmente na busca textual dinâmica, exportação de dados e relatórios, filtros adicionais e orientações de pesquisa.

Em resumo, o MVP proposto e *StelaExperta*<sup>®</sup> oferecem um conjunto mais completo de funcionalidades, enquanto o Portal Integra e Portal Universidade 360 possuem características específicas que podem ser úteis, dependendo das necessidades dos usuários. Para escolher a plataforma mais aderente às demandas das ICTs e da sociedade, é importante avaliar quais funcionalidades são mais relevantes para as necessidades específicas de cada usuário e considerar fatores como custo, facilidade de uso e suporte oferecido pelos fornecedores das plataformas.

A análise das soluções tecnológicas e científicas permitiu identificar possíveis *gaps* não contemplados ou sugeridos. Nesse sentido, o MVP proposto apresenta oportunamente soluções que:

- i) agreguem os dados de PD&S em uma só plataforma;
- ii) permitam o acesso aberto, amplo, transparente e auditável dos resultados de PD&S das ICTs;
- iii) sejam um canal de interação que apresente uma interface amigável, a qual potencialize a transferência, o licenciamento e a solicitação de serviços relacionados às soluções científicas e ou tecnológicas;
- iv) tornem transparente a captação e aplicação dos investimentos públicos e/ou privados direcionados ao desenvolvimento científico e ou tecnológico;
- v) permitam comparar os resultados de PD&S no tempo. As comparações podem ser realizadas entre instituições e/ou entre agrupamentos resultantes dos canais PD&S, que foram organizadas no painel geral chamado PD&S ou nos demais painéis, representando cada atividade

específica de PD&S e considerando o nível de granularidade dos dados atualmente disponíveis;

- vi) apresentem métricas e KPIs mais aderentes aos indicadores utilizados pelo Banco Mundial e pelos governos para estabelecer as políticas para o setor, por exemplo: colaboração Universidade-Indústria em Pesquisa & Desenvolvimento, qualidade das instituições de pesquisa científica, despesas em pesquisa e desenvolvimento (% do PIB), fomento dos SPP em PD&S, dentre outros.

O MVP proposto possui 3 módulos que podem ser disponibilizados no formato *online* e *offline*. O primeiro exhibe os resultados amplos referentes aos canais de PD&S investigadas e alguns KPIs gerais. O segundo apresenta essas atividades em detalhes e alguns KPIs específicos associados a elas. O terceiro permite uma análise simultânea e comparativa de, no mínimo, duas (2) métricas distintas.

Os módulos **Amplio**, **Detalhado** ou **Comparativo** permitem várias análises, considerando, ao menos, catorze (14) canais/categorias/atividades principais de interação (D'este *et al.*, 2007) de PD&S ou, ainda, de suas 24 métricas.

No módulo **Amplio**, ao se analisar o gráfico de barras que traz o número de consultorias prestadas pela RFEPCT por setor TH, é possível constatar que há elevada diferença numérica entre atender as demandas de outros ecossistemas e as dos SPP. Outra constatação é que, apesar da grande diferença quantitativa e da diferença no formato da curva de Consultorias por Ano, pode-se perceber que o ano de 2014 foi o momento com maior número de consultorias, seja com interação com o setor produtivo ou não.

A análise pode trazer também o questionamento sobre que tipo de consultoria foi realizada, e o MVP proposto permite, por exemplo, o acesso a alguns filtros específicos, ampliando a possibilidade analítica por parte do usuário. Por exemplo, ao se desejar saber informações pormenorizadas sobre um determinado número de consultorias em um determinado ano, ao clicar na nuvem de palavras sobre um termo específico, são apresentados os títulos das consultorias vinculados a tal termo e vice-versa.

A modelagem de BI proposta permite que o usuário possa identificar o que tem sido realizado, quem mais realiza e até se realmente o que está descrito como consultoria deveria estar neste grupo. O MVP proposto traz ainda a opção de realizar diversos cruzamentos de dados, como, por exemplo, o de titulação máxima dos pesquisadores, a fim de se visualizar se a titulação interfere no número de consultorias, projetos e orientações de pesquisa concluídas, em andamento ou desativadas. Além disso, apresenta

uma área para pergunta e resposta inteligente e um mapa de densidade dos dados baseado em filtros.

Ainda no que diz respeito à definição geográfica dos dados, o MVP proposto permite a aplicação dos resultados em mapas estaduais, regionais ou nacionais, demonstrando a concentração, ou não, de determinadas ações, trazendo uma visualização similar à apresentada por Johnson *et al.* (2022) ou por Azevedo *et al.* (2021).

Em particular, a ferramenta proposta por Azevedo *et al.* (2021) apresenta quatro módulos (Pesquisa, Ensino, Extensão e Gestão), dos quais apenas Pesquisa foi considerado, para essa análise comparativa com o MVP proposto. Eles criaram um conjunto de KPIs em que não foram apresentados exemplos, sugerindo confidencialidade. Por fim, eles propuseram uma solução de *dashboard* para o IFMT que atendia exclusivamente suas unidades. Comparativamente, o MVP proposto como produto dessa tese permite, para além de outras possibilidades, efetuar a comparação entre instituições, unidades da federação e regiões, podendo ser expandido para comparações entre países em versão futura, além de ser um produto/serviço que abrange todas as ICTs da RFEPCT.

Essas e outras possibilidades, que estão presente no MVP proposto, contribuem para tomadas de decisão e propostas de políticas e ações para fomentar a inovação e a TH que sejam mais assertivas e baseadas em dados (Subseção 4.3). Por exemplo, ao serem analisadas as “orientações de pesquisas” direcionadas a atender aos possíveis problemas dos SPP citados, foi possível constatar que a RFEPCT apresentou uma taxa média de crescimento ano a ano de 45,77%, no período de 2008 a 2018. Sendo que o IFMS apresentou a maior taxa média de crescimento (48,94%) e o IFRR manteve-se sem crescimento no período (0%). Como um dos objetivos e finalidades da RFEPCT é o desenvolvimento nacional, a constatação de tal resultado é um indicador válido, que comprova que a RFEPCT tende a concentrar seus esforços de pesquisa e soluções para atender os setores da economia, os quais o governo brasileiro determina como serviços de maior intensidade de conhecimento, portanto, estratégicos no processo de desenvolvimento inovativo.

Destaca-se ainda que as métricas consideradas neste trabalho tese são compatíveis com estudos no campo da TH, conforme identificado na Seção 2. Os resultados de PD&S da RFEPCT confirmam que suas ICTs tendem a interagir de alguma forma com os SPP, buscando-os por diferentes canais, ainda que inicialmente. Embora os resultados encontrados na literatura não possam ser comparados diretamente, devido aos diferentes períodos, características institucionais e culturais, indiretamente, é possível identificar

tendências e padrões que colaborem para o desenvolvimento de futuras pesquisas, bem como para o planejamento institucional do ecossistema das ICTs brasileiras, em especial, as públicas, as quais uma parte foi aqui analisada.

Por exemplo, no Reino Unido (D'este *et al.*, 2007) entre 2002 e 2003, foram identificados cinco canais/categorias/atividades principais de interação: criação de novas instalações físicas, consultoria e pesquisa por contrato, pesquisa conjunta, treinamento e conferências e reuniões, cada uma refletindo uma ampla gama de interações sem, necessariamente, haver uma sobreposição. De forma similar, foram identificadas seis métricas, neste trabalho, as quais buscam compreender também como essas atividades se desenvolveram no contexto brasileiro, mais especificamente, dentro do ecossistema da RFEPCT.

D'este *et al.* (2007) concluíram que 56% dos pesquisadores universitários em seus estudos se envolveram em consultoria ou pesquisa decorrente da interação com os SPP, pelo menos, uma vez no período de 2002 a 2003. No caso da RFEPCT, no período de 2008 a 2018, o resultado foi de  $\cong 3\%$  (1.233 entre 40.359, pesquisadores). Identificaram também que, para quatro dos cinco tipos de interação investigadas, mais de 40% dos pesquisadores universitários do Reino Unido estiveram envolvidos em alguma forma de interação com os SPP.

Na RFEPCT,  $\cong 5\%$  (1.918 entre 40.359, pesquisadores) estiveram envolvidos em interações com os SPP, em seis das quatorze categorias analisadas. Algumas categorias, como patente, desenho industrial, programa de computador e marca, foram identificadas apenas no nível da ICTs como titular.

Eles também descobriram que a proporção de pesquisadores envolvidos em interações com os SPP variou entre as disciplinas científicas, com maior interação nas disciplinas de Engenharia em comparação com Matemática e Física. Foi possível analisar a tendência para uma distribuição similar na RFEPCT, mesmo sem informações suficientes para serem utilizadas com a integridade e disponibilidade necessária ao nível de granularidade na área de conhecimento.

Considerando, ainda, a categoria “Orientação de Pesquisa” na Rede, houve predominância de pesquisas direcionadas às Ciências Exatas e da Terra  $\cong (22,2\%)$ , seguidas pelas Engenharias  $\cong (17,8\%)$ , Ciências Humanas  $\cong (17,6\%)$  e Ciências Agrárias  $\cong (17,3\%)$ .

Por fim, constatou-se que as interações dos pesquisadores universitários com os parceiros dos SPP, são distribuídas uniformemente pelas regiões do Reino Unido. Na

RFEPCT, relacionada às orientações de pesquisa, as regiões Sudeste, Sul e Centro-Oeste tendem a concentrar o maior número de pesquisas com essa característica.

A análise do módulo **Detalhado** com resultados organizados das informações. No período de 2008 a 2018, seis (6) Instituições de Pesquisa e Ensino (IFES, CEFET/MG, IFCE, IFRN, IFP e IFRS: Institutos Federais do Espírito Santo, Ceará, Rio Grande do Norte, São Paulo e Rio Grande do Sul, e o Centro Federal de Educação Tecnológica de Minas Gerais) foram responsáveis por cerca de 23% das orientações de pesquisa da RFEPCT, somando 284.582, entre 42 instituições analisadas. Essas instituições representam, aproximadamente, 1/4 de todas as orientações de pesquisa em 11 anos, com a UTFPR contribuindo com 17%. Essa observação pode ajudar a entender os fatores que levam a UTFPR a se destacar em alguns estudos.

O MVP também indicou que a maioria das orientações de pesquisa foi conduzida por mestres (52%), seguida de especialistas (27%) e doutores (12%). Aproximadamente, 93% das orientações foram submetidas por pesquisadores com dedicação exclusiva, sugerindo que docentes efetivados em carreiras de magistério superior ou do ensino básico, técnico e tecnológico desempenham papel fundamental no desenvolvimento de pesquisa e inovação nas ICTs brasileiras.

Ao se analisar as orientações de pesquisa em 11 anos, constata-se que cerca de 53% delas envolvem iniciação científica e trabalhos de conclusão de curso. Destes, aproximadamente 33% receberam financiamento em forma de bolsa. Somente 5% dessas bolsas foram financiadas pelos SPP, enquanto cerca de 80% foram financiadas pelas ICTs-GOV, sendo considerado em conjunto devido ao seu caráter público.

A análise do módulo **Comparativo** indica uma queda acentuada no número de orientações de pesquisa de 2017 a 2018, que pode ser atribuída à supressão de recursos públicos para P&D ( $X_1$ ), no mesmo período. Andrade *et al.* (2022) identificaram uma alta correlação  $r(11) = 0,90$ ; valor- $p \leq 0,05$  entre transferência tecnológica e financiamento, sugerindo que a perda de financiamento de P&D pode comprometer a RFEPCT como ecossistema estratégico para o desenvolvimento inovativo.

Esta análise é apenas uma amostra das possibilidades disponíveis, ao considerar todas as ICTs, variáveis, dimensões e subdivisões analisadas. O *dashboard* mostrou-se uma solução viável para apoiar gestores, pesquisadores e outros usuários, atendendo ao princípio básico de apresentar uma interface gráfica amigável, conforme mencionado por Stair *et al.* (2010).

O escopo deste trabalho propõe ferramentas necessárias para manipulação, quantificação e apresentação visual de informações, permitindo uma melhor reflexão e debate em estudos futuros. Novas regras podem ser desenvolvidas e validadas em pesquisas futuras, sendo incorporadas à estrutura de BI proposta com base nas informações quantitativas ou qualitativas disponíveis.

Ao analisar o módulo **Comparativo**, por exemplo pode-se inferir que de 2017 a 2018 o número de *Orientações de Pesquisa* apresentou uma queda acentuada. Mesmo não sendo o objetivo desse trabalho, apenas com uma simples observação da tela do MVP pode-se ter como hipótese, por exemplo, que esse decréscimo foi consequência da supressão de recursos públicos para P&D ( $X_1$ ) no mesmo período. Para a confirmação dessa hipótese é possível, no *dashboard*, realizar um estudo correlacionando tais variáveis, como as disponibilizadas neste módulo, para obter o resultado visual para a análise realizada, baseado em dados. Andrade *et al.* (2022) constaram em um estudo similar, por exemplo que transferência tecnológica e financiamento apresentam uma correlação alta ( $r(11) = 0,90$  (valor- $p \leq 0,05$ )). Esse resultado sugere que caso a Rede perca financiamento de P&D a probabilidade de não transferência tecnológica como uma de suas finalidades e objetivos, o que pode inviabilizar os resultados da RFEPCT como um ecossistema estratégico para o desenvolvimento inovativo.

Ressalta-se que esta é uma pequena demonstração do grande número de possibilidades de análises possíveis de serem extraídas. Se forem consideradas todas as ICTs analisadas, as variáveis de ano, região, unidade da federação e instituições pesquisadas, bem como as 24 métricas gerais e suas subdivisões, é possível efetuar uma enorme variedade de cruzamentos com um enorme número de combinações.

Mesmo com um enorme número de combinações possíveis, o MVP demonstrou ser uma solução viável para suporte não apenas aos gestores, seja como um sistema de informação gerencial, de apoio a decisão e/ou de apoio ao executivo, como também a pesquisadores e demais usuários. O MVP também atende a um dos princípios básicos citados por Stair *et al.* (2010) por apresentar uma interface gráfica mais amigável, não sendo apenas uma série de dados amontoados. Uma vez validadas, novas regras podem ser adicionadas à estrutura de BI, proposta com base em uma ou mais informações geradas a partir das propriedades quantitativas ou qualitativas disponíveis.

### 5.7.2 Estudo de Caso Aplicado: o Impacto das Características Individuais dos Pesquisadores nas Interações entre ICTs-SPP

Mogoutov *et al.* (2008) constataram que as características individuais dos pesquisadores, como idade e antiguidade, têm um impacto muito mais forte na explicação da variedade de interações com a indústria do que as características do departamento ou universidade. Essa descoberta sugere a importância de considerar tais características individuais, ao analisar os canais de transferência de conhecimento.

Por exemplo, ao analisar os resultados *Orientação de pesquisa financiada pelo SPP* ( $X_7$ ) da RFEPCT, foi possível observar que as propostas foram apresentadas por  $\cong$  94% dos pesquisadores com dedicação exclusiva efetivados entre 2 e 10 anos, cerca de 50% dos pesquisadores entre 10,01 e 20 anos, e  $\cong$  44% dos pesquisadores com mais de 20 anos de efetivação. Isso indica que pesquisadores mais jovens e menos experientes tendem a se envolver mais em atividades de transferência de conhecimento, levando-se em conta características individuais como idade e tempo de exercício como efetivo (MVP<sup>32</sup>).

Uma possível explicação para a maior abertura de acadêmicos mais jovens (2,01 a 10 anos) a novas ideias é que eles podem estar no início de suas carreiras, em suas agendas de pesquisa, e ainda não desgastados pelos embates políticos e burocráticos para fomentar um modelo TH de inovação. Além disso, podem ser mais propensos a enxergar o valor das interações com os SPP, mais abertos a correr riscos potenciais e a ter colaboradores de várias áreas do conhecimento, vistos como oportunidades de parceria (Berbegal-Mirabent *et al.*, 2013; D'este *et al.*, 2007).

Ao analisar os resultados da RFEPCT, foi identificada uma forte correlação positiva entre variáveis relacionadas ao ciclo da carreira acadêmica, tais como *Pesquisadores com Dedicação Exclusiva e Tempo de Serviço* entre 2,01 e 10 anos, e a propensão para se envolver em atividades de orientação de pesquisa em colaboração com os SPP. A análise estatística apresenta um valor de  $r(12.184) = 0,46$  e  $p < 0,01$ , indicando a significância dessa correlação. Essa constatação ressalta a importância de se considerar fatores como idade e experiência do pesquisador, ao formular políticas e ações que possam mitigar os desafios enfrentados pelas ICTs, como a evasão de doutores assistentes na Espanha, devido à recessão (Berbegal-Mirabent *et al.*, 2019). Dessa forma, as ICTs

---

<sup>32</sup> <https://prequest.websiteseguro.com/bi>



podem atingir seus objetivos, especialmente os relacionados ao desenvolvimento de produto, processo e/ou serviços, em parceria com os SPP.

Nesse sentido, para evitar um possível êxodo acadêmico futuro, políticas e ações que valorizem e incentivem pesquisadores dedicados são essenciais para a manutenção do desempenho na resolução de problemas conceituais ou aplicados. Alguns exemplos de ações que podem ser implementadas incluem:

- i. Estabelecer programas de mentoria e treinamento para pesquisadores em início de carreira, com foco no desenvolvimento de habilidades de colaboração, comunicação e gestão de projetos em parcerias com SPP;
- ii. Criar políticas de incentivo e reconhecimento para pesquisadores que se envolvem em atividades de transferência de conhecimento, como prêmios, bolsas ou progressão na carreira;
- iii. Fomentar a criação de redes de contato entre pesquisadores, SPP e outros *stakeholders* relevantes, como organizações não governamentais, agências de financiamento e governos, para facilitar a cooperação e o intercâmbio de conhecimentos;
- iv. Desenvolver mecanismos de financiamento específicos para projetos de pesquisa colaborativa entre ICTs e SPP, que possam incentivar o engajamento dos pesquisadores em atividades de inovação e transferência de conhecimento;
- v. Promover a mobilidade de pesquisadores entre ICTs e SPP, por meio de programas de intercâmbio, estágios ou visitas técnicas, para estimular a troca de experiências e a aproximação entre os diferentes atores do ecossistema de inovação;
- vi. Estimular o ensino e a pesquisa aplicada nas ICTs, incluindo temas relevantes para a indústria e a sociedade, incentivando a busca por soluções práticas e inovadoras para problemas reais; e
- vii. Estabelecer parcerias estratégicas entre ICTs e SPP, que possam proporcionar oportunidades de colaboração contínua e de longo prazo, ampliando a capacidade de inovação ou impacto na sociedade.

Ao considerar e implementar tais medidas, as ICTs poderão enfrentar os desafios relacionados à retenção e a motivação de seus pesquisadores, garantindo assim o fortalecimento do ecossistema de inovação e a promoção de uma cultura de cooperação e transferência de conhecimento entre os diferentes atores envolvidos.

Nesta análise, fica evidente o potencial do MVP para identificar precocemente problemas futuros e auxiliar na tomada de decisões. A plataforma se mostra uma opção relevante para detectar possíveis desafios, permitindo a implementação de ações preventivas. No contexto brasileiro, é importante ressaltar que aproximadamente 99% de todas as pesquisas e inovações locais, regionais e nacionais são conduzidas por ICTs públicas (Analytics, 2019). Dessa forma, é crucial desenvolver estratégias que garantam a retenção e motivação dos profissionais que atuam nessas instituições, particularmente, na RFEPCT. O MVP demonstra ser, mais uma vez, uma solução tecnológica promissora nesse processo, considerando a avaliação dos especialistas, contribuindo para a elaboração de políticas efetivas e aprimorando a qualidade da pesquisa e inovação no país.

A análise dos resultados de Berbegal-Mirabent *et al.* (2019) sobre 47 ICTs públicas espanholas revela a falta de evidências de que os Núcleos de Inovação e Tecnologia (NITs ou TTOs, em inglês) estejam contribuindo para a interação efetiva entre ICTs e SPP. No contexto brasileiro, os NITs também são responsáveis por estabelecer essa conexão. Algumas hipóteses futuras podem ser consideradas, como a insegurança jurídica dos pesquisadores e a burocracia na entrega dos resultados esperados pelos SPP. Ao acessar o MVP e analisar o resultado da RFEPCT, será possível obter evidências de que os indicadores relacionados às interações com os SPP tendem a zero após a primeira década.

Com base nessas constatações, algumas diretrizes podem ser propostas em níveis macro, meso e micro da RFEPCT e das ICTs que a compõe para mitigar esse possível problema, como a criação de programas de formação continuada e orientações que conectem os pesquisadores com as demandas dos SPP, bem como treinamento, capacitação e apoio técnico, além de garantir a segurança jurídica para que os pesquisadores possam usufruir dos benefícios da lei nº 13.243 (Brasil, 2016a).

É crucial considerar outros fatores que influenciam a capacidade das ICTs em promover a interação TH. Conforme estabelecido na literatura (Chen *et al.*, 2012), o tamanho e a reputação das instituições são relevantes, pois ICTs maiores e mais conhecidas tendem a ter maior facilidade em se envolver com SPP. Além disso, a quantidade e qualidade da pesquisa produzida são fatores importantes.

Analisando as ICTs da RFEPCT (consultar MVP) com base nesses pressupostos (Chen *et al.*, 2012), observa-se que a UTFPR se destacou em relação às demais. Entre 2008 e 2018, por exemplo, a UTFPR foi responsável por cerca de 44% das patentes

publicadas em parceria com o SPP. No entanto, seu desempenho não foi absoluto, e os resultados da RFEPCT diferem da realidade de Taiwan (Chen *et al.*, 2012).

A literatura enfatiza a importância do financiamento governamental no apoio à transferência de conhecimento e inovação (Dalmarco *et al.*, 2018; Rossi *et al.*, 2015; Zhao *et al.*, 2015). No entanto, há limitações ao depender exclusivamente de fundos públicos para apoiar essas atividades. A precisão e confiabilidade dos dados usados para calcular a distribuição dos fundos e o desenvolvimento tardio das universidades e do setor industrial podem dificultar o estímulo às parcerias entre esses dois grupos (Dalmarco *et al.*, 2018; Rossi *et al.*, 2015).

Apesar desses desafios, o financiamento governamental tem um papel vital no apoio à interação TH, especialmente no caso brasileiro. Considerando que a RFEPCT está em sua fase ainda inicial de evolução, seus gestores podem utilizar as análises e diretrizes apresentadas nesta tese para fomentar o sucesso do modelo TH e contribuir para a elaboração do marco legal que a constitui.

Para maximizar os benefícios do financiamento público, universidades devem focar em projetos de pesquisa aplicada de interesse para parceiros industriais. Isso não implica ignorar a pesquisa básica, mas adotar uma abordagem estratégica para garantir o uso eficaz dos recursos escassos.

Esta análise enfatiza a importância de utilizar o MVP proposto como suporte ao gestor ou a um pesquisador interessado em analisar, por exemplo, as características individuais dos pesquisadores, como idade e tempo de efetivação, ao analisar a transferência de conhecimento entre instituições de ICTs e SPP. Pesquisadores mais jovens e menos experientes tendem a se envolver mais em atividades de transferência de conhecimento, o que pode ser atribuído à sua abertura a novas ideias e menor desgaste causado por embates políticos e burocráticos, como mencionado anteriormente.

Para garantir o sucesso do modelo de interação entre ICTs e SPP, é fundamental desenvolver políticas e ações que valorizem e incentivem pesquisadores dedicados, considerando fatores como idade e experiência. Além disso, é crucial focar em projetos de pesquisa aplicada com potencial interesse para os SPP, garantindo um uso eficaz dos recursos e fortalecimento do ecossistema de inovação.

## 6 Conclusões

O desenvolvimento da solução tecnológica proposta, baseou-se na técnica de *Business Intelligence* (BI), fundamenta nas teorias de Análise de Negócio, Visualização de Dados e da Tripla Hélice de Inovação, para ecossistema das ICTs. Revela-se, segundo os especialistas, como uma tecnologia fundamental, inovadora e importante para a consolidação de informações sobre a performance dessas instituições, relacionadas aos canais de PD&S.

A proposta de um MVP (Produto Mínimo Viável) surge como uma solução inovadora, preenchendo uma lacuna identificada na literatura e, na prática, pelos especialistas que o avaliaram. A teoria da Tripla Hélice (TH) demonstra ser um marco teórico relevante, elucidando a colaboração entre ICTs, setor privado produtivo (SPP) e governo (GOV), conforme ilustrado em capítulos anteriores, em especial para o caso aplicado da RFEPCT.

Esta tese propôs uma análise multidimensional das capacidades inovadoras das ICTs brasileiras. Utilizando-se de técnicas estatísticas como Análise por Componentes Principais (PCA), Análise Hierárquica (HC) e Análise Não Hierárquica (NHC), buscou delinear o modelo Tripla Hélice (TH). A pesquisa também apresentou uma solução tecnológica *online* para monitorar os resultados dos canais de PD&S das ICTs, com potencial de atualização em tempo real, na versão de produção.

A investigação abrangeu uma análise longitudinal das ICTs vinculadas à RFEPCT onde foi possível delinear o modelo TH. A proposta incluiu um modelo analítico integrado baseado em BI, incorporando análise multivariada para gerar *insights* sobre os canais de PD&S e sua interação com o modelo TH de inovação.

A ferramenta tecnológica proposta permite ao usuário personalizar sua análise conforme o nível de detalhamento desejado. Esta solução atende a uma demanda dos gestores das ICTs, fornecendo uma ferramenta *online* em BI que responde rapidamente às exigências dos órgãos de controle brasileiros.

A plataforma, aqui proposta, foi planejada para se integrar a outras, sem a concorrência de dados e informações, como, por exemplo: a NITData, desenvolvida por um grupo de pesquisa da Universidade Estadual de Santa Cruz, que exibe informações quantitativas da Rede NIT-NE, em um cartograma; o INFOVIS, do portal de aplicativos interativos de visualização de dados da Bahia; e com a plataforma Nilo Peçanha,

vinculada ao Ministério da Educação, entre outras soluções locais, regionais ou nacionais existentes.

A tese também contribui ao introduzir a proposta do Índice Integrado de Performance e Inovação em Pesquisa, Desenvolvimento Tecnológico e Prestação de Serviços Técnicos Especializados (IIP\_PD&S). Este índice oferece uma avaliação objetiva do desempenho das ICTs em relação ao SNI, identificando padrões e relações entre elas e fornecendo *insights* sobre sua taxionomia e contribuição ao modelo TH.

O IIP\_PD&S destaca-se como uma contribuição essencial deste trabalho, com vistas à resolução de problemas, tais como o papel das ICTs no ecossistema de inovação brasileiro e orientação de políticas públicas e decisões baseadas em dados.

A tese também contribui a comparação da Rede Federal de Educação Profissional, Científica e Tecnológica do Brasil (RFEPCT) com redes internacionais evidencia desafios significativos para o país. A implementação de tecnologias, como BI, *dashboards* e *Data Warehouse*, destaca-se como uma necessidade para otimizar a gestão e tomada de decisões nas ICTs.

Outra contribuição foi a contatação, também, que embora a RFEPCT seja heterogênea, em algumas situações ela deve adotar padrões de entendimento e aplicação, comuns em todas as regiões, tais como políticas, diretrizes e ações fomentadas em nível macro.

Os resultados obtidos na análise dos canais de PD&S na RFEPCT indicam uma diversidade nas estratégias e desempenhos das ICTs. Algumas instituições se destacam em áreas específicas, enquanto outras apresentam oportunidades para melhorias. A influência do contexto geográfico e regulatório no desempenho das ICTs é evidente, com regiões como o Sudeste e Sul do Brasil apresentando maior inovação e desempenho. O Nordeste também se destaca, mas com apenas instituições como IFCE e IFPE, sobressaindo-se em projetos colaborativos e consultoria para SPP, respectivamente.

A adoção de ferramentas tecnológicas, como o MVP proposto, aliada à capacitação e investimento em áreas estratégicas, é crucial para potencializar o impacto das ICTs no cenário global de inovação, ao permitir também a possibilidade de estudos comparativos *cross-country* e nacionais. A colaboração entre ICTs, SPP e GOV, sob a perspectiva da teoria da TH, é fundamental para o desenvolvimento socioeconômico do Brasil.

## 7 Propostas para Futuros Trabalhos

Ao longo desta tese, explorou-se diversos aspectos relacionados ao tema central, abordando suas nuances, desafios e possibilidades. Constatou-se que, apesar dos avanços significativos alcançados até o momento, ainda há um vasto campo de estudo a ser explorado e aprimorado, com o potencial de gerar contribuições valiosas para a área.

Neste sentido, são propostos uma série de trabalhos futuros que visam expandir o escopo da pesquisa atual e explorar novos caminhos no desenvolvimento e aplicação do conhecimento adquirido. Os trabalhos futuros sugeridos estão divididos em três categorias principais: aprofundamento das questões teóricas, avanços metodológicos e aplicações práticas.

No que diz respeito às questões teóricas, propõem-se investigar ainda mais as inter-relações entre os conceitos-chave discutidos nesta tese, buscando compreender melhor como eles se complementam e se influenciam mutuamente. Além disso, sugere-se a realização de estudos comparativos entre diferentes contextos e realidades, a fim de identificar padrões e particularidades que possam enriquecer a compreensão geral do tema.

No âmbito metodológico, acredita-se que há espaço para aprimorar as abordagens utilizadas, bem como explorar novas técnicas e ferramentas que possam contribuir para uma análise mais precisa e abrangente dos dados e fenômenos investigados em três subcampos de investigações, com o intuito de triangulação dos resultados e validação das conclusões alcançadas.

O primeiro trata da análise multivariada de dados, cujo objetivo é explorar as relações complexas entre múltiplas variáveis em um conjunto de dados. Ela é usada para identificar padrões, detectar relacionamentos entre variáveis, fazer previsões e resumir informações importantes em grandes conjuntos de dados com várias dimensões.

A análise multivariada permite que os pesquisadores testem hipóteses e avaliem a importância relativa de cada variável em relação às outras. Ela é amplamente utilizada em muitos campos, incluindo a Psicologia, Ciências Sociais, Biologia, Economia, Finanças, Marketing, entre outros.

A análise multivariada envolve o uso de técnicas estatísticas avançadas, como análise de componentes principais, análise fatorial, análise discriminante, análise de *clusters*, regressão múltipla, entre outras, para explorar as interações complexas entre as variáveis em um conjunto de dados.

O segundo trata da análise de texto, conhecido como "mineração de texto" (*Text mining*) ou "processamento de linguagem natural" cujo objetivo é explorar e extrair informações úteis e *insights* de grandes volumes de dados textuais, sejam eles em formato de texto escrito ou fala. Essas técnicas são usadas para descobrir padrões, tendências e relacionamentos em dados não estruturados, que podem ser qualitativos ou quantitativos.

A análise de texto permite que as organizações analisem informações importantes contidas em documentos, *e-mails*, mensagens de texto, comentários em redes sociais, entre outros tipos de dados textuais. Ela pode ser aplicada em muitas áreas, como *marketing*, atendimento ao cliente, análise de risco, pesquisa acadêmica, entre outras.

As técnicas de análise de texto incluem análise de sentimento, classificação de documentos, identificação de tópicos, extração de entidades, resumo automático de texto, entre outras. A aplicação dessas técnicas pode ajudar as organizações a tomar decisões informadas, entender melhor seus clientes e usuários, melhorar a eficiência dos processos empresariais, entre outras vantagens.

O terceiro trata de Processamento Analítico *Online*, cujo objetivo é permitir a análise multidimensional e interativa de grandes conjuntos de dados em tempo real. Através de modelos multidimensionais, os dados são organizados em dimensões, hierarquias e medidas, permitindo uma análise aprofundada de diferentes perspectivas.

O Processamento Analítico *Online* pode ser realizado através de tabelas *pivot*, gráficos, mapas e *dashboards* interativos. Suas informações podem ser facilmente filtradas, agrupadas e hierarquizadas para fornecer uma visão mais completa e profunda dos dados. Essas possibilidades fornecem uma visão aprofundada e completa deles, permitindo que as organizações tomem decisões mais informadas e estratégicas com base em *insights* valiosos.

Quanto às aplicações práticas, vislumbramos um grande potencial na implementação das descobertas e recomendações desta tese em contextos específicos, vinculados ao ecossistema de CT&I e PD&I atendendo demandas de ministérios, fundações de apoio a pesquisa, ICTs e SPP que precisem gerenciar seus passivos e ativos, vinculados aos canais de PD&S.

Por fim, a colaboração contínua entre teoria e prática é essencial para aplicar a proposta deste trabalho de investigação científica e orientar pesquisas futuras diante das evoluções em ciência, tecnologia e inovação, buscando otimizar as contribuições para o desenvolvimento sustentável, o avanço socioeconômico e a interação TH contínua.

## **8 Limitações Dessa Tese**

Conforme descrito em capítulos anteriores, as principais limitações do estudo incluem dificuldades no acesso a dados públicos no Brasil, como *links* quebrados e alterados, falta de padronização entre fontes de dados, diversidade de tipos de dados disponíveis e documentação inadequada, além de procedimentos de segurança e restrições de acesso nem sempre atendendo aos requisitos de confidencialidade. Esses desafios reforçam a necessidade de atualizar o DW com dados recentes e desenvolver estratégias para melhorar a qualidade e acessibilidade dos dados, garantindo a relevância e eficácia da solução no monitoramento e avaliação das políticas de PD&S em ecossistema das ICTs.



## 9 Referências

ABD-ALMAGEED, W.; DAVIS, L. S., 2006, Berlin, Heidelberg. **Density Estimation Using Mixtures of Mixtures of Gaussians**. Springer Berlin Heidelberg. 410-422.

ABNT, A. B. D. N. T. Sistemas de gestão da qualidade - Diretrizes para melhorias de desempenho. **Associação Brasileira de Normas Técnicas**. Disponível em: <http://www.standardconsultoria.com/f/files/bf0e78debcf1ce2087d14749a5e73fd4901297844.pdf>. Acesso em: 15 nov. 2017.

ADEINAT, I. M.; ABDULFATAH, F. H. Organizational culture and knowledge management processes: case study in a public university. **Vine Journal of Information and Knowledge Management Systems**, 49, n. 1, p. 35-53, Mar 2019. Article. 10.1108/vjikms-05-2018-0041.

AHMED MUAYAD, Y.; MUSLIM NAJEEB, Z. *et al.* The strategic impact of business intelligence in terms of essentials, techniques, and services. **European Scholar Journal**, 3, n. 5, p. 26-37, 05/14 2022.

AIN, N.; VAIA, G. *et al.* Two decades of research on business intelligence system adoption, utilization and success – A systematic literature review. **Decision Support Systems**, 125, 2019. Article. 10.1016/j.dss.2019.113113.

ALBATS, E.; ALEXANDER, A. T. *et al.* Traditional, virtual, and digital intermediaries in university-industry collaboration: exploring institutional logics and bounded rationality. **Technological Forecasting and Social Change**, 177, p. 22, Apr 2022. Article. 10.1016/j.techfore.2022.121470.

ALBATS, E.; FIEGENBAUM, I. *et al.* A micro level study of university industry collaborative lifecycle key performance indicators. **The Journal of Technology Transfer**, 43, n. 2, p. 389-431, 2018/04/01 2018. <https://doi.org/10.1007/s10961-017-9555-2>.

AMMAR, M. B.; AYACHI, F. L. *et al.* Data warehouse for machine learning: application to breast cancer diagnosis. **Procedia Computer Science**, 196, p. 692-698, 2022/01/01/2022. <https://doi.org/10.1016/j.procs.2021.12.065>.

ANALYTICS, C. Research in Brazil: A report for CAPES by Clarivate Analytics. [recurso eletrônico]. Disponível em: <http://www.capes.gov.br/images/stories/download/diversos/17012018-CAPES-InCitesReport-Final.pdf>.

ANALYTICS, C. Research in Brazil: Funding excellence: Analysis prepared on behalf of CAPES by the Web of Science Group. [recurso eletrônico].

ANDRADE, E. P.; PEREIRA, J. D. S. *et al.* An exploratory analysis of Brazilian universities in the technological innovation process. **Technological Forecasting and Social Change**. Disponível em: <https://doi.org/10.1016/j.techfore.2022.121876>. Acesso em: 22 Nov 2022.

ANKRAH, S.; AL-TABBAA, O. Universities-industry collaboration: A systematic review. **Scandinavian Journal of Management**, 31, n. 3, p. 387-408, Sep 2015. Review. 10.1016/j.scaman.2015.02.003.

ARCURI, M. Políticas de CT&I e financiamento público à infraestrutura de C&T: comparações internacionais e mapeamento da infraestrutura nacional. **Sistemas setoriais de inovação e infraestrutura de pesquisa no Brasil**, p. 581-615, 2016.

AZEVEDO, A.; AZEVEDO, J. M. *et al.*, 2021, **Designing and Implementing a Dashboard with Key Performance Indicators for a Higher Education Institution**. 165-172.

AZEVEDO, S. U. D. **Modelagem do public value scorecard como instrumento de avaliação de desempenho para uma organização do terceiro setor**. 2012. (Mestrado) - Ciências Sociais Aplicadas, Universidade Federal do Paraná, Curitiba, PR

BAGLIERI, D.; BALDI, F. *et al.* University technology transfer office business models: One size does not fit all. **Technovation**, 76-77, p. 51-63, 2018. 10.1016/j.technovation.2018.05.003.

BALAKRISHNAN, G. **Business Intelligence Solution**. 2022. (Mestrado) - Information and Communication Technology, Metropolia University of Applied Sciences, Helsínquia, Finlândia.

BANK'S, W. GovData360. [recurso eletrônico]. **World Bank's**, Pennsylvania Ave. NW, Washington, D.C. Disponível em: <https://govdata360.worldbank.org/>.

BARAHAMA, A.; WARDANI, R., 2021, **Utilization Extract, Transform, Load For Developing Data Warehouse In Education Using Pentaho Data Integration**. IOP Publishing. 012030. Disponível em: <https://iopscience.iop.org/article/10.1088/1742-6596/2111/1/012030/pdf>.

BARRAGAN-OCANA, A.; REYES-RUIZ, G. *et al.* Approach to the identification of an alternative technological innovation index. **Scientometrics**, 122, n. 1, p. 23-45, Jan 2020. Article. 10.1007/s11192-019-03292-9.

BARTLETT, M. S. A note on the multiplying factors for various  $\chi^2$  approximations. **Journal of the Royal Statistical Society. Series B (Methodological)**, p. 296-298, 1954.

BASSIL, Y. A data warehouse design for a typical university information system. **Journal of Computer Science & Research**, 1, n. 6, p. 12-17, 2012. <https://doi.org/10.48550/arXiv.1212.2071>.

BASSO, F. G.; PEREIRA, C. G. *et al.* Cooperation and technological areas in the state universities of São Paulo: An analysis from the perspective of the triple helix model. **Technology in Society**, 65, 2021. Article. 10.1016/j.techsoc.2021.101566.

BEKKERS, R.; FREITAS, I. M. B. Analysing preferences for knowledge transfer channels between universities and industry: To what degree do sectors also matter? **Research policy**, 37, n. 10, p. 1837-1853, 2008.

BENTO, L. M. **As transformações sócio-históricas da Rede Federal de Educação Tecnológica sob o ponto de inflexão informacional: um olhar sobre a produção seriada do conhecimento institucional**. 2015. (Mestrado) - Programa de Pós Graduação em Ciência da Informação, Universidade Federal do Rio de Janeiro, Rio de Janeiro, RJ.

BERBEGAL-MIRABENT, J.; GIL-DOMÉNECH, D. *et al.* Dealing with heterogeneity: an analysis of spanish universities. **Tec Empresarial**, 13, n. 3, p. 58-77, 2019.

BERBEGAL-MIRABENT, J.; LAFUENTE, E. *et al.* The pursuit of knowledge transfer activities: An efficiency analysis of Spanish universities. **Journal of Business Research**, 66, n. 10, p. 2051-2059, 2013. Article. 10.1016/j.jbusres.2013.02.031.

BERNARDINO, C. F.; DEBORTOLI, J. V. *et al.* Triple Helix analysis from Fapemig data for the last 10 years. **Innovation & Management Review**, 17, n. 4, p. 431-446, Dec 2020. Article. 10.1108/inmr-11-2019-0141.

BERTAGNOLLI, F. Key Performance Indicators. *In: Lean Management*: Springer, 2022. p. 307-317.

BIANCHI, I. S.; VENDRUSCOLO, J. D. B. G. *et al.* Business intelligence e dashboards na educação superior: uma revisão sistemática da literatura. **Encontro Internacional de Gestão, Desenvolvimento e Inovação (EIGEDIN)**, 6, n. 1, 2022.

BILLOR, N.; HADI, A. S. *et al.* BACON: blocked adaptive computationally efficient outlier nominators. **Computational Statistics & Data Analysis**, 34, n. 3, p. 279-298, 2000/09/28/ 2000. [https://doi.org/10.1016/S0167-9473\(99\)00101-2](https://doi.org/10.1016/S0167-9473(99)00101-2).

BIMONTE, S.; BILLAUD, O. *et al.* Collect and analysis of agro-biodiversity data in a participative context: A business intelligence framework. **Ecological Informatics**, 61, 2021. <https://doi.org/10.1016/j.ecoinf.2021.101231>.

BOVAIRD, T. Beyond engagement and participation: User and community coproduction of public services. **Public Administration Review**, 67, n. 5, p. 846-860, Sep-Oct 2007. Article. 10.1111/j.1540-6210.2007.00773.x.

BRASIL. Decreto nº 3175, de 2 de dezembro de 1837. Convertendo o Seminário de S. Joaquim em collegio de instrução secundaria, com a denominação de Collegio de Pedro II, e outras disposições. **Coleção de Leis do Império do Brasil**, Rio de Janeiro, RJ, pp. 59.

BRASIL. Decreto nº 7566, de 23 de setembro de 1909. Crêa nas capitaes dos Estados da República Escolas de Aprendizes Artífices, para o ensino profissional primário e gratuito. **Diário Oficial da União**, Brasília, DF.

BRASIL. Lei nº 378, de 13 de janeiro de 1937. Transforma as escolas de aprendizes e artífices mantidas pela União em liceus industriais e institui novos liceus, com o objetivo de propagar nacionalmente o ensino profissional em todos os ramos e graus. **Diário Oficial da União**, Brasília, DF, pp. 1.

BRASIL. Decreto-Lei nº 4.048, de 22 de janeiro de 1942. Cria o Serviço Nacional de Aprendizagem dos Industriários (SENAI). **Diário Oficial da União**, Brasília, DF.

BRASIL. Decreto-Lei nº 4.073, de 30 de janeiro de 1942. Lei orgânica do ensino industrial. **Diário Oficial da União**, Brasília, DF, pp. 1997.

BRASIL. Decreto-Lei nº 4.127, de 25 de fevereiro de 1942. Estabelece as bases de organização da rede federal de estabelecimentos de ensino industrial. **Diário Oficial da União**, Brasília, DF, pp. 2957.

BRASIL. Lei no 3.552, de 16 de fevereiro de 1959. Dispõe sobre nova organização escolar e administrativa dos estabelecimentos de ensino industrial do Ministério da Educação e Cultura, e dá outras providências. **Diário Oficial da União**, Brasília, DF, pp. 1.

BRASIL. Lei nº 5.540, de 28 de novembro de 1968. Fixa normas de organização e funcionamento do ensino superior e sua articulação com a escola média, e dá outras providências. **Diário Oficial da União**, Brasília, DF, pp.

BRASIL. Lei no 6.545, de 30 de junho de 1978. Dispõe sobre a transformação das Escolas Técnicas Federais de Minas Gerais, do Paraná e Celso Suckow da Fonseca em Centros Federais de Educação Tecnológica e dá outras providências. **Diário Oficial da União**, Brasília, DF, pp. 1.

BRASIL. Lei nº 8.948, de 8 de dezembro de 1994. Dispõe sobre a instituição do Sistema Nacional de Educação Tecnológica e dá outras providências. **Diário Oficial da União**, Brasília, DF, pp. 1.

BRASIL. Lei nº 11.184, de 7 de outubro de 2005. Dispõe sobre a transformação do Centro Federal de Educação Tecnológica do Paraná em Universidade Tecnológica Federal do Paraná e dá outras providências. **Diário Oficial da União**, Brasília, DF, pp. 1.

BRASIL. Lei nº 11.892, de 29 de dezembro de 2008. Institui a Rede Federal de Educação Profissional, Científica e Tecnológica, cria os Institutos Federais de Educação, Ciência e Tecnologia, e dá outras providências. **Diário Oficial da União**, Brasília, DF, pp.

BRASIL. Lei nº 12.527, de 18 de novembro de 2011. Regula o acesso a informações previsto no inciso XXXIII do art. 5º, no inciso II do § 3º do art. 37 e no § 2º do art. 216 da Constituição Federal; altera a Lei nº 8.112, de 11 de dezembro de 1990; revoga a Lei nº 11.111, de 5 de maio de 2005, e dispositivos da Lei nº 8.159, de 8 de janeiro de 1991; e dá outras providências. **Diário Oficial da União**, Brasília, DF, 18 nov. 2011, pp. 1.

BRASIL. Lei nº 13.243, de 11 de janeiro de 2016. Dispõe sobre estímulos ao desenvolvimento científico, à pesquisa, à capacitação científica e tecnológica e à inovação e altera a Lei no 10.973, de 2 de dezembro de 2004, a Lei no 6.815, de 19 de agosto de 1980, a Lei no 8.666, de 21 de junho de 1993, a Lei no 12.462, de 4 de agosto de 2011, a Lei no 8.745, de 9 de dezembro de 1993, a Lei no 8.958, de 20 de dezembro de 1994, a Lei no 8.010, de 29 de março de 1990, a Lei no 8.032, de 12 de abril de 1990, e a Lei no 12.772, de 28 de dezembro de 2012, nos termos da Emenda Constitucional no 85, de 26 de fevereiro de 2015. **Presidência da República/Casa Civil/Subchefia para Assuntos Jurídicos**, Brasília, DF, pp. 1.

BRASIL. Mundo Afora - Educação Profissional e Tecnológica. [recurso eletrônico]. **Ministério das Relações Exteriores/Subsecretaria-geral de cooperação internacional, promoção comercial e temas culturais/Departamento Cultural**, Brasília, DF. Disponível em: [https://sistemas.mre.gov.br/kitweb/datafiles/Oslo/pt-br/file/09\\_Cultural/09-10-Mundo\\_Afora\\_14.pdf](https://sistemas.mre.gov.br/kitweb/datafiles/Oslo/pt-br/file/09_Cultural/09-10-Mundo_Afora_14.pdf). Acesso em: Nov, 2016.

BRASIL. Nota Técnica nº 124. Expansão da Rede Federal de Educação Profissional, Científica e Tecnológica quadriênio 2011-2014. [recurso eletrônico]. **Ministério da Educação - Secretaria de Educação Profissional e Tecnológica**, Brasília, DF. Disponível em: <http://portal.mec.gov.br/docman/maio-2016-pdf/41241-nota-tecnica-124-2015-pdf/file>. Acesso em: Jun, 2022.

BRASIL. Lei Nº 13.709. Dispõe sobre a proteção de dados pessoais e altera a Lei nº 12.965, de 23 de abril de 2014 (Marco Civil da Internet). **Presidência da República/Casa Civil/Subchefia para Assuntos Jurídicos**, Brasília, DF, pp. 1.

BRASIL. **Plataforma Nilo Peçanha**. 2018b. Disponível em: <https://www.gov.br/mec/pt-br/pnp>. Acesso em: jan. 07, 2021

BRASIL. Licitações e contratos: orientações e jurisprudência do Tribunal de Contas da União. pp.

BRASIL. Rede Federal completa 110 anos com 661 escolas no país. [recurso eletrônico]. **Ministério da Educação**, Brasília, DF. Disponível em: <http://portal.mec.gov.br/component/content/article/209-noticias/564834057/78121-rede-federal-completa-110-anos-com-661-escolas-no-pais?Itemid=164>. Acesso em: Fev, 2020.

BRASIL. Observatório de Ciência, Tecnologia e Inovação (OCTI). [recurso eletrônico]. **Centro de Gestão e Estudos Estratégicos, Ministério da Ciência, Tecnologia e Inovações (MCTI)**, Brasília, DF. Disponível em: <https://octi.cgee.org.br/sobre>. Acesso em: Ago, 2021.

BRASIL. Painel Universidade 360. **Ministério da Educação**. Disponível em: <https://www.gov.br/mec/pt-br/universidade360/painel-universidade-360>.

BRASIL. Portaria MCTI N° 5.109. Define as prioridades, no âmbito do Ministério da Ciência, Tecnologia e Inovações, no que se refere a projetos de pesquisa, de desenvolvimento de tecnologias e inovações, para o período 2021 a 2023. **16 ago. 2021**, pp.

BRASIL. Relatório de Avaliação: Rede Federal de Educação Profissional, Científica e Tecnológica, Ciclo 2021. [recurso eletrônico]. **CGU - Controladoria-Geral da União**, Brasília, DF. Disponível em: <https://www.gov.br/economia/pt-br/acesso-a-informacao/participacao-social/conselhos-e-orgaos-colegiados/cmmap/politicas/2021/gastos-diretos/rfepct-relatorio-de-avaliacao.pdf>. Acesso em: Nov, 2021.

BRASIL. Lista de Alto Risco da Administração Pública Federal. [recurso eletrônico]. **Tribunal de contas da União (TCU)**, Brasília, DF. Disponível em: [https://portal.tcu.gov.br/data/files/1E/07/C0/FC/925628102DFE0FF7F18818A8/lista\\_de\\_alto\\_risco\\_da\\_administracao\\_publica.pdf](https://portal.tcu.gov.br/data/files/1E/07/C0/FC/925628102DFE0FF7F18818A8/lista_de_alto_risco_da_administracao_publica.pdf). Acesso em: Set, 2022a.

BRASIL. Plataforma web Nilo Peçanha. **Ministério da Educação - Secretaria de Educação Profissional e Tecnológica**, Brasília, DF. Disponível em: <https://www.gov.br/mec/pt-br/pnp>. Acesso em: Jan, 2023.

BREAKSPEAR, A. A new definition of intelligence. **Intelligence and National Security**, 28, n. 5, p. 678-693, 2013.

BRESCHI, S.; MALERBA, F. *et al.* Systems of innovation: technologies, institutions and organizations. *In: Systems of innovation: technologies, institutions and organizations*, 1997.

BROWN, C. E. Coefficient of variation. *In: Applied multivariate statistics in geohydrology and related sciences*: Springer, 1998. p. 155-157.

CABELKOVA, I.; NORMANN, R. *et al.* The Role of Higher Education Institutions in Fostering Industry Clusters in Peripheral Regions: Strategies, Actors and Outcomes. **Higher Education Policy**, 30, n. 4, p. 481-498, Dec 2017. Article. 10.1057/s41307-017-0059-3.

ČÁBELKOVÁ, I.; NORMANN, R. *et al.* The Role of Higher Education Institutions in Fostering Industry Clusters in Peripheral Regions: Strategies, Actors and Outcomes. **Higher Education Policy**, 30, n. 4, p. 481-498, 2017. Article. 10.1057/s41307-017-0059-3.

CAHYADI, A.; PRANANTO, A. Reflecting design thinking: A case study of the process of designing dashboards. **Journal of Systems and Information Technology**, 17, n. 3, p. 286-306, 2015.

CAI, Y. Z.; ETZKOWITZ, H. Theorizing the Triple Helix model: Past, present, and future. **Triple Helix**, 7, n. 2-3, p. 189-226, Jun 2021. Article. 10.1163/21971927-bja10003.

CARAYANNIS, E. G.; CAMPBELL, D. F. 'Mode 3'and'Quadruple Helix': toward a 21st century fractal innovation ecosystem. **International journal of technology management**, 46, n. 3-4, p. 201-234, 2009. doi:10.1504/ijtm.2009.023374.

CARAYANNIS, E. G.; CAMPBELL, D. F. J. Triple Helix, Quadruple Helix and Quintuple Helix and How Do Knowledge, Innovation and the Environment Relate To Each Other? **International Journal of Social Ecology and Sustainable Development**, 1, n. 1, p. 41-69, 2010. 10.4018/jsesd.2010010105.

CARAYANNIS, E. G.; CAMPBELL, D. F. J. *et al.* Helix Trilogy: the Triple, Quadruple, and Quintuple Innovation Helices from a Theory, Policy, and Practice Set of Perspectives. **Journal of the Knowledge Economy**, 13, n. 3, p. 2272-2301, 2022/09/01 2022. 10.1007/s13132-021-00813-x.

CARTAXO, R. M.; GODINHO, M. M. How institutional nature and available resources determine the performance of technology transfer offices. **Industry and Innovation**, 24, n. 7, p. 713-734, 2017. 10.1080/13662716.2016.1264068.

CASTILLO, W.; MEDINA QUISPE, F. *et al.* Una metodología para procesos data warehousing basada en la experiencia. 2018.

CESARONI, F.; PICCALUGA, A. The activities of university knowledge transfer offices: towards the third mission in Italy. **The Journal of Technology Transfer**, 41, n. 4, p. 753-777, 2015. 10.1007/s10961-015-9401-3.

CESARONI, F.; PICCALUGA, A. The activities of university knowledge transfer offices: towards the third mission in Italy. **Journal of Technology Transfer**, 41, n. 4, p. 753-777, 2016. Article. 10.1007/s10961-015-9401-3.

CHASIN, F., 2016, **Business analysis of digital discourse for new service development: a theoretical perspective and a method for uncovering the structure of social representations for improved service development**. IEEE. 1567-1576.

CHAUDHURI, S.; DAYAL, U. An overview of data warehousing and OLAP technology. **ACM Sigmod record**, 26, n. 1, p. 65-74, 1997.

CHAVES, A. O.; DA CRUZ, G. P. Eficiência Do Instituto Federal Baiano: Análise Dos Grupos De Pesquisa E Propriedade Industrial. **Cadernos de Prospecção**, 10, n. 3, 2017. 10.9771/cp.v10i3.23120.

CHEN, S.-H.; HUANG, M.-H. *et al.* Driving factors of external funding and funding effects on academic innovation performance in university–industry–government linkages. **Scientometrics**, 94, n. 3, p. 1077-1098, 2012. 10.1007/s11192-012-0864-9.

CHESBROUGH, H. The logic of open innovation: Managing intellectual property. **California Management Review**, 45, n. 3, p. 33+, Spr 2003. Article. 10.2307/41166175.

COHEN, J. **Statistical Power Analysis for the Behavioral Sciences**. 2 ed. NJ: Lawrence Erlbaum.: Hillsdale, 1988. 1483276481.

CORBUCCI, P. R. Dimensões estratégicas e limites do papel da educação para o desenvolvimento brasileiro. **Revista Brasileira de Educação**, 16, n. 48, p. 563-584, 2011.

D'ESTE, P.; PATEL, P. University-industry linkages in the UK: What are the factors underlying the variety of interactions with industry? **Research Policy**, 36, n. 9, p. 1295-1313, 2007. Article. 10.1016/j.respol.2007.05.002.

DA CUNHA, L.; FERREIRA, M. S. *et al.* The effect of long-term climatic variability on wild mammal populations in a tropical forest hotspot: A business intelligence framework. **Ecological Informatics**, 73, p. 101924, 2023.

DALMARCO, G.; SILVEIRA, L. *et al.* Knowledge flow in university-industry relations: an analysis of companies established at a technology park in brazil. 2018.



DAMYANOV, I.; TSANKOV, N. On the Possibilities of Applying Dashboards in the Educational System. **Tem Journal-Technology Education Management Informatics**, 8, n. 2, p. 424-429, May 2019. Article. 10.18421/tem82-15.

DASHDONDOV, K.; SONG, M.-H. Factorial Analysis for Gas Leakage Risk Predictions from a Vehicle-Based Methane Survey. **Applied Sciences**, 12, n. 1, p. 115, 2022.

DE CARVALHO, B. G.; TONELLI, D. F. Limites e Possibilidades do Marco Legal da CT&I de 2016 para as Instituições Científicas e Tecnológicas do Brasil. **Revista de Administração, Sociedade e Inovação**, 6, n. 2, p. 6-24, 2020. doi.org/10.20401/rasi.6.2.356.

DE NEGRI, F.; SQUEFF, F. D. H. S. O mapeamento da infraestrutura científica e tecnológica no brasil. 2016.

DE NEGRI, F. O.; SQUEFF, F. D. H. S. O. Sistemas setoriais de inovação e infraestrutura de pesquisa no Brasil. *In*, 2016.

DENWATTANA, N.; SAENGSAI, A. *et al.*, 2016, English, Maejo Univ, Chiang Mai, THAILAND. **A Framework of Thailand Higher Education Dashboard System**. NEW YORK: Ieee, 2016. Disponível em: <https://ieeexplore.ieee.org/document/7859883/>.

DEVLIN, B. **Data warehouse: from architecture to implementation**. Addison-Wesley Longman Publishing Co., Inc., 1996. 0201964252.

DI TRIA, F.; LEFONS, E. *et al.* Academic Data Warehouse Design Using a Hybrid Methodology. **Computer Science and Information Systems**, 12, n. 1, p. 135-160, Jan 2015. Article. <https://doi.org/10.2298/csis140325087d>.

DIOUF, P. S.; BOLY, A. *et al.* Variety of data in the ETL processes in the cloud: State of the art. **2018 IEEE International Conference on Innovative Research and Development (ICIRD)**, p. 1-5, 11-12 May 2018. <https://doi.org/10.1109/ICIRD.2018.8376308>.

DORE, R. Technology policy and economic-performance - lessons from Japan - FREEMAN, C. **Research Policy**, 17, n. 5, p. 309-310, Oct 1988. Book Review. 10.1016/0048-7333(88)90011-x.

DRUCKER, P. F. **Post-Capitalist Society** (New York: Harper Business, 1993). **Chapter**, 6, p. 123, 1993.

DURO NOVOA, V.; PEREZ CUEVAS, C. M. Business Intelligence System and Decision Support of Economic Management at the University of Havana. **3C TIC**, 5, n. 4, p. 38-54, 2016.

ECKERSON, W.; WHITE, C. Evaluating ETL and data integration platforms. *In: Report of The Data Warehousing Institute*, 2003. v. 184.

EL-GAMMAL, A.; ABDELLATIF, T. *et al.* Using Online Visual Merchandising to Enhance Web Usability (A study on E-government websites in Egypt). **Journal of Design Sciences and Applied Arts**, 3, n. 1, p. 90-96, 2022.

ERDOS, F.; IEEE, 2019, English, Naples, ITALY. **Economical Aspects of UX Design and Development**. NEW YORK: Ieee, 2019. 211-214. Disponível em: <https://ieeexplore.ieee.org/document/9089992/>.

ESTER, M.; KRIEGEL, H.-P. *et al.*, 1996, **A density-based algorithm for discovering clusters in large spatial databases with noise**. 226-231.

ETZKOWITZ, H. The triple helix: Academic-industry-government relations - Implications for the New York regional innovation environment. *In: RAYMOND, S. U. (Ed.). Technology Link to Economic Development: Past Lessons and Future Imperatives*. New York: New York Acad Sciences, 1996. v. 787, p. 67-86. (Annals of the New York Academy of Sciences).

ETZKOWITZ, H.; LEYDESDORFF, L. The Triple Helix--University-industry-government relations: A laboratory for knowledge based economic development. **EASST review**, 14, n. 1, p. 14-19, 1995.

ETZKOWITZ, H.; LEYDESDORFF, L. The dynamics of innovation: from National Systems and "Mode 2" to a Triple Helix of university-industry-government relations. **Research Policy**, 29, n. 2, p. 109-123, Feb 2000. Article. 10.1016/s0048-7333(99)00055-4.

EVERITT, B. S.; LANDAU, S. *et al.* Hierarchical clustering. **Cluster analysis**, 5, p. 71-110, 2011.

FAN, X.; YANG, X. W. *et al.* Diversified resources and academic influence: patterns of university-industry collaboration in Chinese research-oriented universities. **Scientometrics**, 104, n. 2, p. 489-509, Aug 2015. Article. 10.1007/s11192-015-1618-2.

FANA, W. S.; PERMANA, R. *et al.* Data Warehouse Design With ETL Method (Extract, Transform, And Load) for Company Information Centre. **International Journal of Artificial Intelligence Research**, 5, n. 2, p. 132-137, 2021. <https://doi.org/10.29099/ijair.v5i2.215>.

FARINA, M.; PEREZ, C. *et al.* **Psicodinâmica das cores em comunicação**. Editora Blucher, 2011. 8521216939.

FÁVERO, L. P.; BELFIORE, P. **Manual de análise de dados: estatística e modelagem multivariada com Excel®, SPSS® e Stata®**. Elsevier Brasil, 2017. 8535285059.

FÁVERO, L. P.; BELFIORE, P. **Manual de análise de dados: estatística e modelagem multivariada com Excel®, SPSS® e Stata®**. Elsevier Brasil, 2020. 8535285059.

FERNANDEZ, M.; DAVILA, A. *et al.*, 2017, Spanish, Cartagena, COLOMBIA. **Data quality applied to an academic business intelligence solution: lesson learned**. NEW YORK: Ieee, 2017. Disponível em: [Go to ISI://WOS:000419845200013](https://doi.org/10.1109/WOS:000419845200013)

<https://ieeexplore.ieee.org/document/8088200/>.

FIELD, A. **Discovering statistics using IBM SPSS statistics: North American edition**. sage, 2017. 152644030X.

FIELD, A. **Discovering Statistics Using IBM SPSS Statistics**. SAGE Publications, 2018. 9781526419521.

FIGUEIREDO, N.; FERNANDES, C. Cooperation University-Industry: A Systematic Literature Review. **International Journal of Innovation and Technology Management**, 17, n. 08, p. 36, Dec 2020. Review. 10.1142/s0219877021300019.

FIKRI, N.; RIDA, M. *et al.* An adaptive and real-time based architecture for financial data integration. **Journal of Big Data**, 6, n. 1, p. 1-25, 2019.

FREEMAN, C. **Technology, policy, and economic performance: lessons from Japan**. Burns & Oates, 1987. 0861879287.

FU, L., 2016, **A Recommendation System Using OLAP Approach**. IEEE. 622-625.

GALATI, F.; BIGLIARDI, B. *et al.* Why do academics become entrepreneurs? How do their motivations evolve? Results from an empirical study. **International Journal of Entrepreneurial Behavior & Research**, 26, n. 7, p. 1477-1503, Oct 2020. Article. 10.1108/ijebr-11-2019-0619.

GALVÃO, A.; MASCARENHAS, C. *et al.* Triple helix and its evolution: a systematic literature review. **Journal of Science and Technology Policy Management**, 10, n. 3, p. 812-833, Oct 2019. Review. 10.1108/jstpm-10-2018-0103.

GARCIA-ALSINA, M.; ESPINET, E. O. Inteligencia competitiva: corpus teórico y prácticas. **Ibersid: revista de sistemas de información y documentación**, 6, p. 77-88, 2012.

GARCIA, J. C. R. **Indicadores de produção dos grupos de pesquisa do Instituto Federal de Educação, Ciência e Tecnologia da Paraíba**. 2011. (Mestrado) - Ciência da Informação, Universidade Federal da Paraíba, João Pessoa, PB.

GARSON, G. D. Cluster analysis. **Asheboro, NC: Statistical Associates Publishers**, 2014.

GAVRILOVA, T. A.; ALSUFYEV, A. I. *et al.* Knowledge visualization: critique of the St. Gallen School and an analysis of contemporary trends. **Business Informatics**, n. 3, p. 7-19, 2017.

GIROTTO, E. D.; DE OLIVEIRA, J. V. P. *et al.* Gestão e transparência de dados educacionais na rede estadual paulista (1995-2018): a hegemonia gerencial. **Revista Brasileira de Política e Administração da Educação**, 36, n. 3, p. 1248-1272, 11/12 2020. 10.21573/vol36n32020.99902.

GOMES, M. C.; PIMENTEL, T. R. *et al.* GO PISA! Indicativos para elaboração de políticas públicas educacionais. **Revista Educação e Políticas em Debate**, 8, n. 2, p. 233-253, 2019.

GRAY, D. E. **Pesquisa no mundo real**. Penso Editora, 2016. 8563899295.

GUITART, I.; CONESA, J., 2015, English, Taipei, TAIWAN. **Analytic Information Systems in the context of Higher Education: Expectations, reality and trends**. NEW YORK: Ieee, 2015. 294-300. Disponível em: [<Go to ISI>://WOS:000380529500048](https://ieeexplore.ieee.org/document/7312087/)  
<https://ieeexplore.ieee.org/document/7312087/>.

HAIR, J. F.; BLACK, W. C. *et al.* **Multivariate data analysis**. Prentice hall Upper Saddle River, NJ, 2019.

HAMPEL, G. Preparing the Conceptual Model of a Database. **QUAESTUS MULTIDISCIPLINARY RESEARCH JOURNAL**, 19, n. Jun, p. 290-304, 2021.

HAN, J.; KAMBER, M. *et al.* Data mining concepts and techniques third edition. **University of Illinois at Urbana-Champaign Micheline Kamber Jian Pei Simon Fraser University**, 2012.

HARTIGAN, J. A. **Clustering algorithms**. John Wiley & Sons, Inc., 1975. 047135645X.

HASSAN, B. A.; QADER, S. M. A New Framework to Adopt Multidimensional Databases for Organizational Information System Strategies. 2021. <https://doi.org/10.48550/arXiv.2105.08131>.

HEANG, R. Literature review of business intelligence. 2017.

HEUKELOM, F. Measurement and Decision Making at the University of Michigan in the 1950s and 1960s. **Journal of the History of the Behavioral Sciences**, 46, n. 2, p. 189-207, 2010.

HIDALGO, A.; ALBORS, J. Innovation management techniques and tools: a review from theory and practice. **R & D Management**, 38, n. 2, p. 113-127, Mar 2008. Article. 10.1111/j.1467-9310.2008.00503.x.

HOFFMANN, R. T. **Desenvolvimento de um dashboard para auxiliar a gestão de pessoas em uma universidade federal**. 2021. (Mestrado profissional) - Administração, Universidade Federal de Santa Catarina, Florianópolis, SC. Disponível em: <https://repositorio.ufsc.br/handle/123456789/229269>.

HOTELLING, H. Analysis of a complex of statistical variables into principal components. **Journal of educational psychology**, 24, n. 6, p. 417, 1933.

HUBERT, L.; ARABIE, P. Comparing partitions. **Journal of classification**, 2, p. 193-218, 1985.

IBGE. Estimativas de população. [recurso eletrônico]. **Instituto Brasileiro de Geografia e Estatística**, Brasília, DF. Disponível em: [ftp://ftp.ibge.gov.br/Estimativas\\_de\\_Populacao/Estimativas\\_2019/estimativa\\_dou\\_2019.pdf](ftp://ftp.ibge.gov.br/Estimativas_de_Populacao/Estimativas_2019/estimativa_dou_2019.pdf). Acesso em: Jun, 2020.

IBGE. Estatísticas. [recurso eletrônico]. **Instituto Brasileiro de Geografia e Estatística**, Brasília, DF. Disponível em: <https://www.ibge.gov.br/estatisticas/downloads-estatisticas.html>. Acesso em: Out, 2022.

INMON, W. H. What is a data warehouse. **Prism Tech Topic**, 1, n. 1, p. 1-5, 1995.

IQBAL, M. Z.; MUSTAFA, G. *et al.*, 2020, Singapore. **A Review of Star Schema and Snowflakes Schema**. Springer Singapore. 129-140. Disponível em: [https://link.springer.com/chapter/10.1007/978-981-15-5232-8\\_12](https://link.springer.com/chapter/10.1007/978-981-15-5232-8_12).

JACKSON, P.; MAVI, R. K. *et al.* University-industry collaboration within the triple helix of innovation: The importance of mutuality. **Science and Public Policy**, 45, n. 4, p. 553-564, Aug 2018. Article. 10.1093/scipol/scx083.

JAMIU, S. M.; ABDULLAH, N. S. *et al.*, 2020, English, **Data governance support for business intelligence in higher education: A systematic literature review**. Springer. 35-44. Disponível em: [https://www.scopus.com/inward/record.uri?eid=2-s2.0-85077769499&doi=10.1007%2f978-3-030-33582-3\\_4&partnerID=40&md5=eaa2797472b43514f6f1597806934750](https://www.scopus.com/inward/record.uri?eid=2-s2.0-85077769499&doi=10.1007%2f978-3-030-33582-3_4&partnerID=40&md5=eaa2797472b43514f6f1597806934750).

JAPAL, A.; SUTEDI, S., 2021, **Designing a Lecturer's Performance Data Warehouse Model Using Star Scheme**. 73-76. Disponível em: <https://jurnal.darmajaya.ac.id/index.php/icitb/article/view/3046>.

JOHNSON, E.; HEMMATIAN, I. *et al.* A Framework and Databases for Measuring Entrepreneurial Ecosystems. **Research Policy**, 51, n. 2, p. 9, Mar 2022. Article. <https://doi.org/10.1016/j.respol.2021.104398>.

JOVANOVIĆ, M.; SAVIĆ, G. *et al.* Towards a Triple Helix based efficiency index of innovation systems. **Scientometrics**, 127, n. 5, p. 2577-2609, 2022/05/01 2022. 10.1007/s11192-022-04304-x.

KAISER, H. F. A second generation little jiffy. **Psychometrika**, 35, n. 4, p. 401-415, 1970.

KANG, W. M.; ZHAO, S. L. *et al.* Triple helix in the science and technology innovation centers of China from the perspective of mutual information: a comparative study between Beijing and Shanghai. **Scientometrics**, 118, n. 3, p. 921-940, Mar 2019. Article. 10.1007/s11192-019-03017-y.

KAPLAN, R. S. **The balanced scorecard measures that drive performance**. Harvard business review, 1992. 71-79 p.

KAPLAN, R. S.; NORTON, D. P. **A estratégia em ação: balanced scorecard**. 22 ed. Rio de Janeiro: Campus: Gulf Professional Publishing, 1997. 344 p. 8535201491.

KASHEFI, A.; ALWZINANI, F. *et al.*, 2018, **Perspectives on teaching modeling and simulation in a department of computer science**. IEEE. 4058-4068.

KHATIBI, V.; KERAMATI, A. *et al.* A Business Intelligence Approach to Monitoring and Trend Analysis of National R&D Indicators. **Engineering Management Journal**, 29, n. 4, p. 244-257, 2017/10/02 2017. <https://doi.org/10.1080/10429247.2017.1380578>.

KIM, J.-O.; AHTOLA, O. *et al.* **Introduction to factor analysis: What it is and how to do it**. Sage, 1978. v. 13). 0803911653.

KIM, J.; DAIM, T. U. *et al.* University technology transfer: A conceptual model of impacting factors and phased process. *In: PICMET '09 - 2009 Portland International Conference on Management of Engineering & Technology*, 2009, 2-6 Aug. 2009. p. 2798-2811. DOI: 10.1109/PICMET.2009.5261803. Disponível em: <https://ieeexplore.ieee.org/document/5261803/>.

KIMBALL, R.; ROSS, M. *et al.* **The Data Warehouse Lifecycle Toolkit, 2nd Edition**. January 2008. 978-0-470-14977-5. 672 p. Disponível em: <https://www.kimballgroup.com/data-warehouse-business-intelligence-resources/books/data-warehouse-dw-lifecycle-toolkit/>.

KIYANI, A. A.; AYUPP, K. *et al.* Construct validation of job characteristics in an asian context: an evidence among academicians in private universities in Pakistan. **International Journal of Business and Society**, 19, n. 3, p. 816-832, 2018. Article.

KONGTHANASUWAN, T.; SRIWIBOON, N. *et al.* Market Analysis with Business Intelligence System for Marketing Planning. **Information**, 14, n. 2, p. 116, 2023.

KRMAC, E. V. Intelligent Value Chain Networks: Business Intelligence and Other ICT Tools and Technologies in Supply/Demand Chains. *In: Supply/Demand Chains. Supply Chain Management - New Perspectives*, 2011.

KUMAR, S.; KUMAR, N. *et al.* Detection of Brain Tumor Region in MRI Image Through K-Means Clustering Algorithms. *In: Trends and Advancements of Image Processing and Its Applications*: Springer, 2022. p. 221-232.

LAURSEN, K.; REICHSTEIN, T. *et al.* Exploring the Effect of Geographical Proximity and University Quality on University-Industry Collaboration in the United Kingdom. **Regional Studies**, 45, n. 4, p. 507-523, 2011. Article. 10.1080/00343400903401618.

LENARDUZZI, V. and Taibi, D. 2016. Mvp explained: A systematic mapping study on the definitions of minimal viable product. *Software Engineering and Advanced Applications (SEAA)*, 2016 42th Euromicro Conference on (2016), 112– 119.

LEYDESDORFF, L. The Triple Helix, Quadruple Helix, ..., and an N-Tuple of Helices: Explanatory Models for Analyzing the Knowledge-Based Economy? **Journal of the Knowledge Economy**, 3, n. 1, p. 25-35, 2012/03/01 2012. 10.1007/s13132-011-0049-4.

LEYDESDORFF, L. N-Tuple of helices, 2013, Citeseer, p. 1400-1402.

LEYDESDORFF, L.; DOLFSMA, W. *et al.* Measuring the knowledge base of an economy in terms of triple-helix relations among 'technology, organization, and territory'. **Research Policy**, 35, n. 2, p. 181-199, Mar 2006. Article. 10.1016/j.respol.2005.09.001.

LEYDESDORFF, L.; ETZKOWITZ, H. Emergence of a Triple Helix of university-industry-government relations. **Science and Public Policy**, 23, n. 5, p. 279-286, 1996. Article.

LEYDESDORFF, L.; ETZKOWITZ, H. Triple Helix of innovation: Introduction. **Science and Public Policy**, 25, n. 6, p. 358-364, 1998. Article.

LEYDESDORFF, L.; PORTO-GOMEZ, I. Measuring the expected synergy in Spanish regional and national systems of innovation. **Journal of Technology Transfer**, 44, n. 1, p. 189-209, Feb 2019. Article. 10.1007/s10961-017-9618-4.

LEYDESDORFF, L.; SUN, Y. National and International Dimensions of the Triple Helix in Japan: University-Industry-Government Versus International Coauthorship Relations. **Journal of the American Society for Information Science and Technology**, 60, n. 4, p. 778-788, Apr 2009. Article. 10.1002/asi.20997.

LI, M.; HE, L. *et al.* The triple helix system and regional entrepreneurship in China. **Entrepreneurship & Regional Development**, 32, n. 7-8, p. 508-530, 2019. 10.1080/08985626.2019.1666168.

LI, Y.; TANG, Y. J. A dynamic capabilities perspective on pro-market reforms and university technology transfer in a transition economy. **Technovation**, 103, p. 15, May 2021. Article. 10.1016/j.technovation.2021.102224.

LIAO, H.-T.; WANG, Z. *et al.* Developing a Minimum Viable Product for Big Data and AI Education: Action Research Based on a Two-Year Reform of an Undergraduate Program of Internet and New Media. *In: Proceedings of the 2019 4th International Conference on Big Data and Computing, 2019, Guangzhou, China. Association for Computing Machinery*, p. 42-47. DOI: <https://doi.org/10.1145/3335484.3335509>. Disponível em: <https://doi.org/10.1145/3335484.3335509>.

LINO, S. R. L. **Diretrizes para a institucionalização da gestão do conhecimento na rede federal de educação profissional, científica e tecnológica, Brasil**. 2013. (Doutorado) - Engenharia e Gestão do Conhecimento Universidade Federal de Santa Catarina, Florianópolis, SC. Disponível em: <https://repositorio.ufsc.br/bitstream/handle/123456789/123126/325459.pdf?sequence=1&isAllowed=y>.

LIPKOWITZ, I. **Business Economics**, 4, n. 4, p. 89-91, 1969.

LIU, K., 2020, **The Cultivation of Business Analysis Talents in the Background of Big Data**. Atlantis Press. 1-5.



LLOYD, S. Least squares quantization in PCM. **IEEE transactions on information theory**, 28, n. 2, p. 129-137, 1982.

MACEDO, T. L. D. S. **Composição de Grupos de Pesquisa e sua Relação com Redes de Coautorias: O caso do IFBA**. 2014. (Mestrado) - Administração, Universidade Federal da Bahia, Salvador, BA. Disponível em: <https://repositorio.ufba.br/ri/bitstream/ri/24053/1/Thais%20Leite%20da%20Silva%20Macedo.pdf>. Acesso em: Novembro 12, 2017.

MACHADO, V. B. F., 2007, Campinas, SP. **Reformas da educação profissional: a “modernização do arcaico”**. Faculdade de Educação, UNICAMP: Centro Federal de Educação Tecnológica de Minas Gerais - CEFET/MG. Disponível em: [www.histedbr.fe.unicamp.br/acer\\_histedbr/encontro/encontro1/trab\\_pdf/t\\_valeria%20boognini.pdf](http://www.histedbr.fe.unicamp.br/acer_histedbr/encontro/encontro1/trab_pdf/t_valeria%20boognini.pdf). Acesso em: Jan, 2020.

MACQUEEN, J., 1967, **Classification and analysis of multivariate observations**. University of California Los Angeles LA USA. 281-297.

MACQUEEN, J., 1967, **Some methods for classification and analysis of multivariate observations**. Oakland, CA, USA. 281-297.

MARCHEZAN, M. L.; PACHECO, R. C. D. S. *et al.*, 2012, **Gestão estratégica de informações curriculares em ICTIs**.

MARQUES, Y. B.; BAX, M. P. *et al.* Dados abertos dos Institutos Federais de Educação: análise de temas sugeridos pela Controladoria-Geral da União. **AtoZ: novas práticas em informação e conhecimento**, 11, p. 1-10, 2022.

MARTINS, N.; MARTINS, S. *et al.* Design principles in the development of dashboards for business management. **Perspectives on Design II: Research, Education and Practice**, p. 353-365, 2022.

MCGILL, W. J. Multivariate information transmission. **Psychometrika**, 19, n. 2, p. 97-116, Jun 1954. Article.

MEDINA, Q. F.; FARIÑA M., F. *et al.* Data Mart para obtención de indicadores de productividad académica en una universidad. **Ingeniare. Revista chilena de ingeniería**, 26, p. 88-101, 2018.

MENEZES, C. C. N.; PAIXÃO, P. B. S. *et al.* Mapeamento da propriedade intelectual e da transferência de tecnologia do instituto federal de sergipe: um estudo para intensificar as estratégias de proteção e transferência de tecnologia. 2019.

METCALFE, S. The economic foundations of technology policy: equilibrium and evolutionary perspectives. **Handbook of the economics of innovation and technological change**, 1995.

MIANO, V. Y.; COUTO, C. L. P. D. *et al.* Grupos de Pesquisa dos Institutos Federais no RJ – crescimento do eixo Pesquisa e suas áreas. **Revista de Ciência e Inovação**, 2, n. 2, p. 57-67, 2017. 10.26669/2448-4091165.

MIANO, V. Y.; DO COUTO, C. L. P. *et al.* Atipicidades em grupos de pesquisa dos Institutos Federais do Rio de Janeiro: levantamento dos quantitativos em 2018 de acordo os critérios da CNPq. **Revista Gestão em Análise**, 9, n. 1, p. 64-73, 2020.

MIANO, V. Y.; DO COUTO, C. L. P. *et al.* Grupos de Pesquisa dos Institutos Federais no RJ – crescimento do eixo Pesquisa e suas áreas. **Revista de Ciência e Inovação**, 2, n. 2, p. 57-67, 2017.

MILLS, C. W. The structure of power in American society. **The British Journal of Sociology**, 9, n. 1, p. 29-41, 1958.

MIRANDA, J. L. D. S.; FELIX, F. D. N. *et al.* Análise de Agrupamentos: uma aplicação aos Institutos Federais de Educação e Tecnologia em 2016. *In*: Congresso Norte-Nordeste de Pesquisa e Inovação, 2018, Recife.

MLS, K.; GAVALEC, M., 2009, **Multi-criteria models in autonomous decision making systems**. 1-8.

MOGOUTOV, A.; CAMBROSIO, A. *et al.* Biomedical innovation at the laboratory, clinical and commercial interface: A new method for mapping research projects, publications and patents in the field of microarrays. **Journal of Informetrics**, 2, n. 4, p. 341-353, 2008. 10.1016/j.joi.2008.06.005.

MOULD, D.; D'HAENS, G. *et al.* Clinical decision support tools: the evolution of a revolution. **Clinical Pharmacology & Therapeutics**, 99, n. 4, p. 405-418, 2016.

MURTAGH, F.; LEGENDRE, P. Ward's Hierarchical Agglomerative Clustering Method: Which Algorithms Implement Ward's Criterion? **Journal of Classification**, 31, n. 3, p. 274-295, 2014. 10.1007/s00357-014-9161-z.

MUSA, S.; ALI, N. B. M. *et al.*, 2019, English, **Success factors for business intelligence systems implementation in higher education institutions – A review**. Springer Verlag. 322-330. Disponível em: [https://www.scopus.com/inward/record.uri?eid=2-s2.0-85053896412&doi=10.1007%2f978-3-319-99007-1\\_31&partnerID=40&md5=801bea2b949e3521d04a3605fd9d7e06](https://www.scopus.com/inward/record.uri?eid=2-s2.0-85053896412&doi=10.1007%2f978-3-319-99007-1_31&partnerID=40&md5=801bea2b949e3521d04a3605fd9d7e06).

NALAWADE, S.; JOSHI, A. A Literature Review: Data Mining Techniques, Applications & Issues. **Aayushi International Interdisciplinary Research Journal (AIIRJ)**, VIII, n. IV, 2021.

NASCIMENTO, S. D. F. Nível de colaboração e transferência de conhecimento entre os atores do ecossistema de inovação: proposição de um modelo analítico nos Institutos Federais de Educação, Ciência e Tecnologia. 2021.

NEGRO, A. R.; MESIA, R. The Business Intelligence and Its Influence on Decision Making. **Journal of Applied Business & Economics**, 22, n. 2, 2020.

NELSON, R. R. **National innovation systems: a comparative analysis**. Oxford University Press on Demand, 1993. 0195076176.

NIVEN, P. R. **Balanced Scorecard Passo-a-Passo: Elevando o Desempenho e Mantendo Resultados**. Rio de Janeiro: Qualitymark, 2007. 334 p. (Caderno Economia.

NIWA, T. H. **O modelo da hélice tríplice em consonância com os arranjos produtivos locais nas incubadoras tecnológicas: um estudo de caso nas IUTs da UTFPR**. 2014. (Mestrado) - Planejamento e Governança Pública, Universidade Tecnológica Federal do Paraná, Curitiba, PR. Disponível em: [http://repositorio.utfpr.edu.br/jspui/bitstream/1/1000/1/CT\\_PPGPGP\\_M\\_NIWA%2c%20Tiago%20Hideki\\_2014.pdf](http://repositorio.utfpr.edu.br/jspui/bitstream/1/1000/1/CT_PPGPGP_M_NIWA%2c%20Tiago%20Hideki_2014.pdf). Acesso em: Outubro 13, 2015.

NSANZUMUHIRE, S. U.; GROOT, W. Context perspective on University-Industry Collaboration processes: A systematic review of literature. **Journal of Cleaner Production**, 258, 2020. 10.1016/j.jclepro.2020.120861.

OLIVEIRA, G. S. D.; RENAULT, T. B. A interação com atores da hélice tríplice e as perspectivas de desenvolvimento de capacidades empreendedoras. **Revista de Administração, Sociedade e Inovação**, 6, n. 1, p. 24-42, 2020.

OLSEN, D. 2015. *The lean product playbook: How to innovate with minimum viable products and rapid customer feedback*. John Wiley & Sons.

OZKAYA, G.; TIMOR, M. *et al.* Science, Technology and Innovation Policy Indicators and Comparisons of Countries through a Hybrid Model of Data Mining and MCDM Methods. **Sustainability**, 13, n. 2, p. 694, 2021.

PAIKARAY, B. K.; DUBE, M. R. *et al.* ETL and Business Analytics Correlation Mapping with Software Engineering. *In: Machine Learning and Information Processing*: Springer, 2021. p. 501-511.

PALLONI, R.; FRONTONI, E. *et al.* **Open Data Analytics: Advanced methods, tools and visualizations for policy making**. 2018. 125 f. (Doutorado) - Dipartimento di Ingegneria dell'Informazione, Università Politecnica delle Marche, Ancona, Province of Ancona, Itália.

PARK, H.-S.; JUN, C.-H. A simple and fast algorithm for K-medoids clustering. **Expert Systems with Applications**, 36, n. 2, Part 2, p. 3336-3341, 2009/03/01/ 2009. <https://doi.org/10.1016/j.eswa.2008.01.039>.

PARK, H. W. Transition from the Triple Helix to N-Tuple Helices? An interview with Elias G. Carayannis and David F. J. Campbell. **Scientometrics**, 99, n. 1, p. 203-207, Apr 2014. Editorial Material. 10.1007/s11192-013-1124-3.

PASWAN, J.; SINGH, V. K. *et al.* Does university-industry-government collaboration in research gets higher citation and altmetric impact? A case study from India. **Scientometrics**, 127, n. 11, p. 6063-6082, Nov 2022. Article. 10.1007/s11192-022-04508-1.

PAZ, F.; CAZELLA, S. Academic analytics: a systematic review of literature. **International Journal of Development Research**, 9, p. 31710-31716, 11/11 2019.

PEARSON, K. VII. Mathematical contributions to the theory of evolution.—III. Regression, heredity, and panmixia. **Philosophical Transactions of the Royal Society of London. Series A, containing papers of a mathematical or physical character**, n. 187, p. 253-318, 1896.

PEDROSA, I. Da cor à cor inexistente. *In: Da cor à cor inexistente*, 1995. p. 220-220.

PEIXOTO, M.; BUAINAIN, A. Desempenho e Desafios do Sistema de Propriedade Industrial no Brasil. *In: Brasília: Núcleo de Estudos e Pesquisas/CONLEG/Senado.(Texto para Discussão no 294)*. Disponível em: <http://www.senado.leg.br/estudos>, 2021.

PEREIRA, L. S. **Caracterização da comunidade que utiliza dados abertos governamentais sobre a educação brasileira**. 2022. 115 f. (Mestrado) - Ciência da Computação, Centro de Engenharia Elétrica e Informática, Campina Grande, PB. Disponível em: <http://dspace.sti.ufcg.edu.br:8080/jspui/handle/riufcg/25581>. Acesso em: 2022.

PERERA, I.; MEEDENIYA, D. Enabling SMEs to Use Cloud Computing Services—An Exploratory Study on Enterprises Strategy Alterations. **Publisher: INREWI, Editors: V. Tewari**, 2012.

PERKMANN, M.; TARTARI, V. *et al.* Academic engagement and commercialisation: A review of the literature on university-industry relations. **Research Policy**, 42, n. 2, p. 423-442, Mar 2013. Review. 10.1016/j.respol.2012.09.007.

PERKMANN, M.; WALSH, K. University-industry relationships and open innovation: Towards a research agenda. **International Journal of Management Reviews**, 9, n. 4, p. 259-280, Dec 2007. Review. 10.1111/j.1468-2370.2007.00225.x.

PERUCCHI, V.; MUELLER, S. P. M. Características das atividades de pesquisa dos professores dos Institutos Federais de Educação, Ciência e Tecnologia. **Encontros Bibli: revista eletrônica de biblioteconomia e ciência da informação**, 20, n. 44, 2015. 10.5007/1518-2924.2015v20n44p73.

PERUCCHI, V.; MUELLER, S. P. M. Produção dos professores dos Institutos Federais de Educação, Ciência e Tecnologia no currículo da Plataforma Lattes. **Informação & Informação**, 22, n. 1, 2017. 10.5433/1981-8920.2017v22n1p111.

PORTO-GOMEZ, I.; ZABALA-ITURRIAGAGOITIA, J. M. *et al.* Innovation systems in México: A matter of missing synergies. **Technological Forecasting and Social Change**, 148, 2019. <https://doi.org/10.1016/j.techfore.2019.119721>.

PURBASARI, R.; WIJAYA, C. *et al.* Actor Collaboration in the Entrepreneurial Ecosystem: Triple Helix Approach. **Review of Integrative Business and Economics Research**, 9, p. 157-175, 2020.

QUEVEDO, T. C.; GERALDI, M. S. *et al.* Applying machine learning to develop energy benchmarking for university buildings in Brazil. **Journal of Building Engineering**, 63, p. 105468, 2023/01/01/ 2023. <https://doi.org/10.1016/j.jobbe.2022.105468>.

RABOVA, I.; ST'ASTNY, J., 2014, English, Sturovo, SLOVAKIA. **Business Process Model of Key Performance Indicators Harvesting and Measuring**. PRAHA 3: Wolters Kluwer Cr a S, 2014. 473-481.

RAMOS-VIELBA, I.; FERNANDEZ-ESQUINAS, M. *et al.* Measuring university-industry collaboration in a regional innovation system. **Scientometrics**, 84, n. 3, p. 649-667, Sep 2010. Article. 10.1007/s11192-009-0113-z.

RAMOS, M. N. Ensino Médio na Rede Federal e nas Redes Estaduais: por que os estudantes alcançam resultados diferentes nas avaliações de larga escala? **Holos**, 2, p. 449-459, 2018. 10.15628/holos.2018.6976.

RANGA, M.; ETZKOWITZ, H. Triple Helix Systems: An Analytical Framework for Innovation Policy and Practice in the Knowledge Society. **Industry and Higher Education**, 27, n. 4, p. 237-262, 2013. Article. 10.5367/ihe.2013.0165.

RANTALA, T.; UKKO, J. Performance measurement in university-industry innovation networks: implementation practices and challenges of industrial organisations. **Journal of Education and Work**, 31, n. 3, p. 247-261, 2018. Article. <https://doi.org/10.1080/13639080.2018.1460655>.

RAPINI, M. S.; CHIARINI, T. *et al.* The intensity of private funding and the results of university? Firm interactions: the case of Brazil. **Innovation & Management Review**, 16, n. 2, p. 161-184, 2019. Article. 10.1108/inmr-11-2018-0088.

RIBEIRO, B. G. G.; ORBOLATO, D. R. S. *et al.* Identificação das áreas de pesquisa no IFTM usando técnica de agrupamento. **Revista Brasileira de Iniciação Científica**, 7, n. 1, p. 149-165, 2020.

RIES, E. **The lean startup: How today's entrepreneurs use continuous innovation to create radically successful businesses.** Currency, 2011. 030788791X.

ROSSI, A. L. **Transferência de tecnologia: um modelo para a Universidade Federal do Rio Grande do Sul.** 2022. (Doutorado) - Faculdade de Ciências Econômicas, Universidade Federal do Rio Grande do Sul, Porto Alegre, RS. Disponível em: <http://hdl.handle.net/10183/248011>.

ROSSI, F.; ROSLI, A. Indicators of university-industry knowledge transfer performance and their implications for universities: evidence from the United Kingdom. **Studies in Higher Education**, 40, n. 10, p. 1970-1991, Nov 2015. Article. 10.1080/03075079.2014.914914.

ROUSSEEUW, P. J. SILHOUETTES - A GRAPHICAL AID TO THE INTERPRETATION AND VALIDATION OF CLUSTER-ANALYSIS. **Journal of Computational and Applied Mathematics**, 20, p. 53-65, Nov 1987. 10.1016/0377-0427(87)90125-7.

SABATO, J.; BOTANA, N. La Ciencia y la Tecnologia em el Desarrollo Futuro de América Latina1: Revista de la Integración, 1968, INTAL.

SAFADI, M.; MA, J. *et al.*, 2016, English, Sydney, AUSTRALIA. **Mapping for the future: Business intelligence tool to map regional housing stock.** AMSTERDAM: Elsevier Science Bv, 2017. 1684-1694. Disponível em: <https://www.sciencedirect.com/science/article/pii/S1877705817318386?via%3Dihub>.

SANTANA, B. S. **Extração e aplicação de indicadores no processo de recomendação de recursos urbanos utilizando dados estruturados e não-estruturados.** 2019. (Mestrado) - Ciência da Computação, Universidade Federal do Rio Grande do Sul, Porto Alegre, RS. Disponível em: <https://www.lume.ufrgs.br/handle/10183/193897>.

SANTI, R. P.; PUTRA, H., 2018, English, **A Systematic Literature Review of Business Intelligence Technology, Contribution and Application for Higher Education**. Institute of Electrical and Electronics Engineers Inc. 404-409. Disponível em: <https://www.scopus.com/inward/record.uri?eid=2-s2.0-85065167810&doi=10.1109%2fICITSI.2018.8696019&partnerID=40&md5=445fa53df8a97e8d7279caffedc9dcc>.

SANTOS, A. P. Institutos federais de educação: fontes de informação e gestão do conhecimento. **Revista ACB**, 15, n. 1, p. 22-38, 2010.

SANTOS, E.; PIMENTEL, C. A. R. *et al.* Aerotaxonomy. **J. Prod. Auto**, 2, p. 41 - 60., 2019.

SANTOS, M. P.; COSTA, L. S. F., 2015, Espanha, Murcia. **Estudo da Produção e Divulgação Científica do Instituto Federal de Educação, Ciência e Tecnologia de São Paulo: Uso do software livre Script Lattes**. Universidad de Murcia.

SANTOS, O.; TIOSSO, F. *et al.* Demonstração dos benefícios do minimum viable product na criação de um novo aplicativo móvel. **Revista Interface Tecnológica**, 16, n. 1, p. 124-135, 06/30 2019.

SANTOS, T. R.; RIBEIRO, M. W. S. *et al.*, 2021, English, **A systematic review on the implementation of business intelligence at federal universities**. IADIS Press. 85-92. Disponível em: <https://www.scopus.com/inward/record.uri?eid=2-s2.0-85124084368&partnerID=40&md5=b76055cdc2f0ceb937f5863f4f948d5f>.

SANTOS, T. R.; RIBEIRO, M. W. S. *et al.* A systematic review on the implementation of business intelligence at federal universities. *In: 2018 International Conference on Information Technology Systems and Innovation*, 2021. p. 85.

SCHINDLER, E. Aproximando Ciência Política e Estatística na era do Big Data. 2016.

ŠEREŠ, L.; PAVLICEVIC, V. *et al.* A Performance Indicators of University-Industry Collaboration. *In: EDULEARN19 Proceedings*, 2019, p. 9664-9672. DOI: 10.21125/edulearn.2019.2413

SHANNON, C. E. A mathematical theory of communication. **Bell System Technical Journal**, 27, n. 4, p. 623-656, 1948. Article. 10.1002/j.1538-7305.1948.tb00917.x.

SHARON, J. A.; JULIET, S., 2022, English, **Efficient Business Intelligence Implementation: A Systematic Review**. Institute of Electrical and Electronics Engineers Inc. 144-149. Disponível em: <https://www.scopus.com/inward/record.uri?eid=2-s2.0->

85133489558&doi=10.1109%2fCAAIC53929.2022.9793012&partnerID=40&md5=df36252faa965b245007b86a597ff4f7.

SIEGEL, D. S.; WALDMAN, D. *et al.* Assessing the impact of organizational practices on the relative productivity of university technology transfer offices: an exploratory study. **Research Policy**, 32, n. 1, p. 27-48, Jan 2003. Article; Proceedings Paper. 10.1016/s0048-7333(01)00196-2.

SIERRA, L. M. P.; VILLAZUL, S. J. J. University-industry linkage and technology transfer management A quantitative study of an entrepreneurial university. **Journal of Management Development**, 37, n. 7, p. 573-585, 2018. Article. 10.1108/jmd-06-2017-0203.

SILVA, S. F. P. E. Fatores propulsores para o desenvolvimento da pesquisa e da inovação na Rede de Educação Profissional, Científica e Tecnológica. 2015.

SINGH, Y. S.; KIRANI, Y. *et al.* An Analytical System: Data Modelling Practices for Handling an Epidemic. *In: Data Management, Analytics and Innovation*: Springer, 2021. p. 447-457.

SINHA, S.; GHOSH, P. *et al.* Employability of fresh engineering graduates in India A fresh look applying Expectation Confirmation Theory. **Education and Training**, 62, n. 1, p. 47-63, Dec 2019. Article. 10.1108/et-12-2018-0265.

SKYRIUS, R. Information Integration. *In: Business Intelligence*: Springer, 2021. p. 81-111.

SMITH, R. C. Jesuit buildings in Brazil. **The Art Bulletin**, 30, n. 3, p. 187-213, 1948. 10.1080/00043079.1948.11407813.

SNEATH, P. H.; SOKAL, R. R. **Principles of Numerical Taxonomy**. San Francisco and London: W. H. Freeman and Company, 1963. 0716706970.

SOARES, M. D. J. A., 1982, **As escolas de aprendizes artífices-estrutura e evolução**. 58-92. Disponível em: <http://bibliotecadigital.fgv.br/ojs/index.php/fe/article/viewFile/60628/58869>. Acesso em: Abr, 2018.

SOKAL, R. R. A statistical method for evaluating systematic relationships. **Univ. Kansas, Sci. Bull.**, 38, p. 1409-1438, 1958.

SOKAL, R. R. Numerical taxonomy. **Scientific American**, 215, n. 6, p. 106-117, 1966.



SOUZANCHI KASHANI, E.; ZARGHAMI, H. R. The dynamics of university-industry-government relationships in Nanoscience: investigating the Triple-Helix differences between Iran and Switzerland. **Technology Analysis & Strategic Management**, 31, n. 7, p. 817-832, 2018. 10.1080/09537325.2018.1554859.

SPEARMAN, C. " General Intelligence" Objectively Determined and Measured. **The American Journal of Psychology**, v.15, n. n.2, p. 201-292, 1904.

SRIVASTAVA, G.; MUNEESWARI, S. *et al.* A review of the state of the art in business intelligence software. **Enterprise Information Systems**, 16, n. 1, p. 1-28, 2022. Review. 10.1080/17517575.2021.1872107.

STAIR, R.; REYNOLDS, G. *et al.* **Principios de sistemas de información.** Cengage Learning Editores, SA De CV, 2010. 6074814449.

STANDARDIZATION, I. O. F.; Technical Committee ISO/IEC JTC 1, i. T. S. S., software *et al.* **Systems and Software Engineering: Systems and Software Quality Requirements and Evaluation (SQuaRE): System and Software Quality Models.** ISO, 2011.

STEINER, M. T. A.; SOMA, N. Y. *et al.* Abordagem de um problema médico por meio do processo de KDD com ênfase à análise exploratória dos dados. **Gestão & Produção**, 13, p. 325-337, 2006.

STEWART, C. L.; DEWAN, M. A. A. A Systemic Mapping Study of Business Intelligence Maturity Models for Higher Education Institutions. **Computers**, 11, n. 11, p. 27, Nov 2022. Review. 10.3390/computers11110153.

SUMNER, M. Knowledge management: Theory and practice. **Proceedings of the ACM SIGCPR Conference**, p. 1-3, 1999. Article.

SUZUKI, H. J. Método educativo para demonstrar que quantidade de pedidos de patentes não é uma boa métrica de inovação. BR 10 2018 004918 6 A2. Brasil 2018.

TAN, A.-H., 1999, **Text mining: The state of the art and the challenges.** 65-70.

TEAM, R. C. **R: A language and environment for statistical computing.** 2022 ed. R Foundation for Statistical Computing, 2022. Disponível em: <https://www.R-project.org/>.

THORNDIKE, R. L. Who belongs in the family? **Psychometrika**, 18, n. 4, p. 267-276, 1953/12/01 1953. 10.1007/BF02289263.

TIJSSEN, R. J. W.; VAN LEEUWEN, T. N. *et al.* Benchmarking university-industry research cooperation worldwide: performance measurements and indicators based on co-authorship data for the world's largest universities. **Research Evaluation**, 18, n. 1, p. 13-24, Mar 2009. Article. 10.3152/095820209x393145.

TINSLEY, H. E.; TINSLEY, D. J. Uses of factor analysis in counseling psychology research. **Journal of counseling psychology**, 34, n. 4, p. 414, 1987. <https://doi.org/10.1037/0022-0167.34.4.414>.

TONG, B.; LI, Y. *et al.* The development and application of China national landslide database and information system. **Arabian Journal of Geosciences**, 14, n. 6, p. 448, 2021/03/06 2021. 10.1007/s12517-021-06825-w.

TRIPATHI, A.; BAGGA, T. *et al.* Strategic impact of business intelligence: A review of literature. **Prabandhan: Indian Journal of Management**, 13, n. 3, p. 35-48, 2020. Review. 10.17010/pijom/2020/v13i3/151175.

TSENG, F. C.; HUANG, M. H. *et al.* Factors of university-industry collaboration affecting university innovation performance. **Journal of Technology Transfer**, 45, n. 2, p. 560-577, Apr 2020. Article. 10.1007/s10961-018-9656-6.

TZORTZAKI, A. M.; MIHIOTIS, A. A review of knowledge management theory and future directions. **Knowledge and Process Management**, 21, n. 1, p. 29-41, 2014.

VILLASENOR, E. A.; ARENCIBIA-JORGE, R. *et al.* Multiparametric characterization of scientometric performance profiles assisted by neural networks: a study of Mexican higher education institutions. **Scientometrics**, 110, n. 1, p. 77-104, Jan 2017. Article. 10.1007/s11192-016-2166-0.

WANG, J. Construction and Implementation of Multi Collaborative Service Platform Based on Industrial Competitive Intelligence. **Open Journal of Applied Sciences**, 13, n. 3, p. 335-342, 2023.

WANG, L.; WANG, G. *et al.* Big Data and Visualization: Methods, Challenges and Technology Progress. **Digital Technologies**, 1, n. 1, p. 33-38, 2015. 10.12691/dt-1-1-7.

WANG, X. G.; JIA, R., 2013, English, Guangzhou, PEOPLES R CHINA. **Design and Implementation of University Information Service Platform Based on Data Warehouse**. DURNTEN-ZURICH: Trans Tech Publications Ltd, 2013. 2543-2550. Disponível em: <https://www.scientific.net/AMM.321-324.2543>.

WANG, Y., 2015, English, **Business intelligence and analytics education: Hermeneutic literature review and future directions in is education**. Americas Conference on Information Systems. Disponível em:

<https://www.scopus.com/inward/record.uri?eid=2-s2.0-84963529400&partnerID=40&md5=ec3a87f0349c2737b042b1295dda703c>.

YAAKUB, M. H.; MOHAMED, Z. A. Measuring the performance of private higher education institutions in Malaysia. **Journal of Applied Research in Higher Education**, 12, n. 3, p. 425-443, Jun 2020. Article. 10.1108/jarhe-10-2018-0208.

YIN, R. K. **Estudo de Caso-: Planejamento e métodos**. Bookman editora, 2015. 8582602324.

ZHAI, D.; HE, W., 2010, **An application of business intelligence based on patent in data integration and analysis**. IEEE. 288-292. Disponível em: <https://ieeexplore.ieee.org/document/5662374/>.

ZHANG, C.; CAO, L. *et al.* On the feature engineering of building energy data mining. **Sustainable Cities and Society**, 39, p. 508-518, 2018/05/01/ 2018. <https://doi.org/10.1016/j.scs.2018.02.016>.

ZHANG, L.; QI, Z. *et al.* A Review on the Construction of Business Intelligence System Based on Unstructured Image Data. **Procedia Computer Science**, 199, p. 392-398, 2022/01/01/ 2022. Disponível em: <https://doi.org/10.1016/j.procs.2022.01.048>.

ZHANG, Z.; MURTAGH, F. *et al.* Hierarchical cluster analysis in clinical research with heterogeneous study population: highlighting its visualization with R. **Annals of translational medicine**, 5, n. 4, p. 75-75, 2017. 10.21037/atm.2017.02.05.

ZHAO, S. L.; CACCIOLATTI, L. *et al.* Regional collaborations and indigenous innovation capabilities in China: A multivariate method for the analysis of regional innovation systems. **Technological Forecasting and Social Change**, 94, p. 202-220, May 2015. Article. 10.1016/j.techfore.2014.09.014.

## 10 APÊNDICES

### **Apêndice A - Histórico da Rede Federal de Educação Profissional, Científica e Tecnológica**

A criação da RFEPCT, embora relativamente recente, resultou de um longo percurso iniciado em 1686. Historicamente, a primeira iniciativa de educação profissional no Brasil começou com a criação do Seminário de Nossa Senhora de Belém perto de Cachoeira no estado da Bahia no ano de 1686 (Smith, 1948). Seu fundador foi o educador jesuíta e padre português Alexandre de Gusmão (1626-1724). Um século e meio depois, em 29 de novembro de 1837, o Colégio Pedro II (CP II) foi fundado pelo imperador Dom Pedro II do Brasil (1825-1891) (Brasil, 1837).

No entanto, considerou-se como marco temporal nesta tese o ano de 1837 como início do que hoje se chama RFEPCT. Tal ano é histórico pois trata-se da criação do CP II, o qual, atualmente, é uma instituição que integra a RFEPCT. A formação da RFEPCT pode ser resumida por alguns marcos históricos ocorridos entre 1837 e 2008, conforme descritos a seguir.

Em 23 de setembro de 1909, o presidente brasileiro Nilo Procópio Peçanha (1867 - 1924) criou 19 Escolas de Aprendizes Artífices voltadas para a educação profissional gratuita, em diferentes regiões do Brasil, sob a jurisdição do Ministério da Agricultura, Indústria e Comércio (Brasil, 1909). O objetivo era a educação de crianças e jovens vulneráveis e socialmente excluídos entre 10 e 13 anos. A implantação ocorreu nas capitais de Alagoas, Amazonas, Bahia, Ceará, Espírito Santo, Goiás, Maranhão, Mato Grosso, Minas Gerais, Paraíba, Pará, Paraná, Piauí, Pernambuco, Rio Grande do Norte, São Paulo, Santa Catarina, Sergipe e também em Campos, cidade do Rio de Janeiro (Soares, 1982).

Em 13 de janeiro de 1937, conforme estabelecido na lei nº 378, artigo 129 o governo transformou as Escolas de Aprendizes Artífices em Escolas Industriais e Técnicas (Brasil, 1937). Ainda, em seu artigo 10º a formação educacional do Brasil foi dividida em 8 subáreas. Dentre essas 8 subáreas, foi criada a subárea de Educação Industrial (Machado, 2007). A criação dessa subárea objetivou atender à crescente demanda industrial por profissionais qualificados.

Em 30 de janeiro de 1942, as Escolas de Aprendizes de Artesãos foram transformadas em Escolas Industriais e Técnicas, oferecendo formação profissional

equivalente ao nível secundário (Brasil, 1942c). Ainda neste ano foi promulgado o decreto lei nº 4073 (Brasil, 1942b), o qual equiparou a formação profissional ao nível secundário. Essa equiparação autorizava que os estudantes formados nesses cursos técnicos pudessem ingressar no ensino superior, como faculdades e universidades, em área equivalente a cursos acadêmicos. Essa transformação estabeleceu as bases para a criação e organização da Associação Nacional de Aprendizagem Industrial e do Serviço Nacional de Formação Industrial (Brasil, 1942b). O objetivo dessas ações foi ampliar e consolidar a formação técnica profissional no país (Brasil, 1942a).

Em 16 de fevereiro de 1959 foram criadas as primeiras Escolas Técnicas Federais (Brasil, 1959). Este ato estabeleceu sua autonomia didática, técnica, financeira e administrativa; um passo necessário para acelerar o processo de industrialização. Durante a década de 1960 foram definidas outras ações, dentre as quais destacam-se a qualificação para oferecer ensino técnico superior, atualmente denominado tecnológico, no Brasil (Brasil, 1968).

Em 30 de junho de 1978, algumas Escolas Técnicas Federais foram transformadas em Centros Federais de Educação Tecnológica (CEFET). Nessa etapa, o CEFET foi qualificado a realizar pesquisas voltadas às demandas industriais e a oferecer cursos de graduação e pós-graduação (Brasil, 1978). Cerca de uma década e meia depois, em 1994, outras Escolas Técnicas Agrícolas também foram transformadas em CEFET por decretos específicos (Brasil, 1994).

Na busca por obrigatoriamente oferecer atividades de ensino, pesquisa e extensão, três CEFET (Paraná, Rio de Janeiro e Minas Gerais) solicitaram sua alteração para o *status* de universidade federal. Entre eles, apenas o CEFET/PR alcançou essa meta, iniciando em 11 de dezembro de 1998 e terminando em 7 de outubro de 2005 (Brasil, 2005), como a primeira universidade do Brasil dedicada ao ensino técnico, denominada Universidade Tecnológica Federal do Paraná.

Em 31 de dezembro de 2008, foi instituída a atual RFEPC e criados os Institutos Federais de Educação, Ciência e Tecnologia, também conhecidos na época como (IF) (ou Institutos Federais), os quais são frutos da junção de Escolas Técnicas Agrícolas, Centros Federais de Educação Tecnológica e Escolas Técnicas Federais (Brasil, 2008). Como consequência, essa regulamentação, indiretamente, lhe atribuiu uma característica considerada *sui generis* (singular) na história da educação profissional no Brasil e quiçá mundial. Tal característica singular pode ser comprovada ao comparar as características

da RFEPCT à estrutura da educação profissional em outros países como Canadá, Alemanha, Espanha, Finlândia, dentre outros<sup>33</sup> (Brasil, 2016b).

Tal característica é especialmente atribuída a RFEPCT, uma vez que representa uma inédita proposta de ensino, pesquisa e extensão pública que oferta, ao mesmo tempo, Cursos de Qualificação - Formação Inicial Continuada (FIC); Educação à Distância (EAD); Ações Afirmativas; Educação de Jovens e Adultos (EJA); Educação Profissional para Jovens e Adultos (PROEJA); Ensino Médio integrado a cursos Técnicos Concomitantes e Subsequentes; Cursos Superiores de Tecnologia, Licenciatura e Bacharelado; e Cursos de Pós-graduação de Especialização (*Lato Sensu*), Mestrado Profissional e Acadêmico e Doutorado (*Stricto Sensu*).

Dentre suas finalidades – lei 11.892/2008 (Brasil, 2008) – consta a formação profissional e tecnológica voltada a atender aos diversos setores produtivos, com ênfase no desenvolvimento socioeconômico local, regional e nacional, orientada a potencializar os arranjos produtivos sociais e culturais locais decorrentes de pesquisa aplicada, da produção cultural, do empreendedorismo, cooperativismo e do desenvolvimento científico e tecnológico, do desenvolvimento e a transferência de tecnologias sociais, preferencialmente relacionadas à preservação do meio ambiente.

Dentre seus objetivos – lei 11.892/2008 (Brasil, 2008) – consta ainda estimular, apoiar, desenvolver e difundir atividades de pesquisa, de desenvolvimento técnico e tecnológico, e de prestação de serviços inerentes ao desenvolvimento regional, preferencialmente. Outra característica, a ser destacada é que os IF, no desenvolvimento da sua ação acadêmica, devem garantir o mínimo de 50% (cinquenta por cento) de suas vagas ao nível de educação médio técnico (preferencialmente integrados) e EJA; e o mínimo de 20% (vinte por cento) para o nível de educação superior (bacharelado, tecnólogo, licenciatura e pós-graduação *stricto e lato sensu*), conforme art. 8º da lei 11.892/2008.

A RFEPCT é constituída por um conjunto de instituições públicas de ensino brasileiras que possuem autonomia administrativa, patrimonial, financeira e didático-pedagógica. Ela consiste atualmente em 2 CEFET com 17 *campi*, 38 IF com 598 *campi*, uma Universidade Tecnológica com 17 *campi*, o CPII com 14 *campi* e as 23 escolas

---

<sup>33</sup> Entretanto, é importante considerar que os contextos e abordagens dessas nações são diversos e não necessariamente podem ser comparados de maneira estrita. Além disso, alguns estudos (Brasil, 2016b) abordam aspectos críticos, como a qualidade do ensino e a empregabilidade dos alunos após a formação, questões essas não abordadas no presente trabalho.

técnicas vinculadas a 14 Universidades Federais (Brasil, 2022b). A Figura A. 1 apresenta o tamanho da RFEPCCT até 31 de dezembro de 2022.

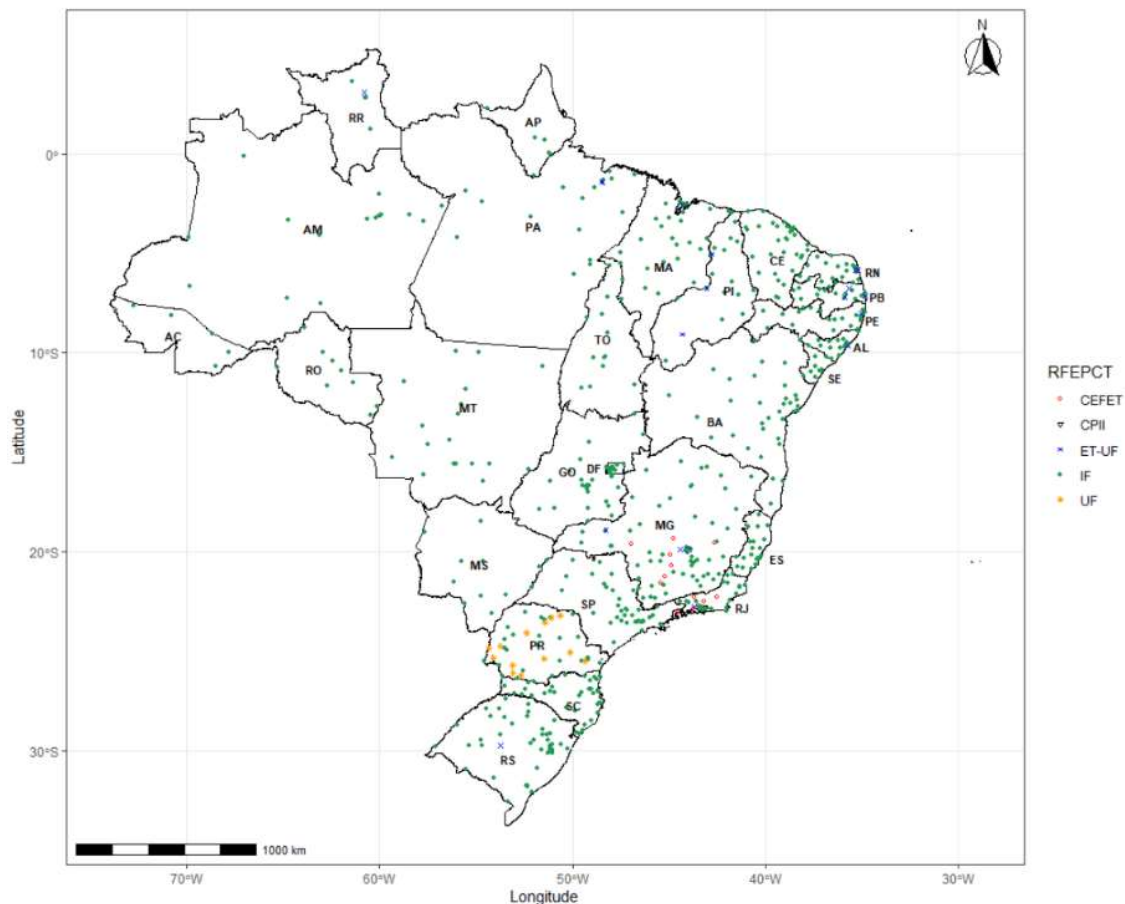


Figura A. 1: Mapa territorial da Rede Federal de Educação Profissional, Científica e Tecnológica demonstrando sua abrangência e capilaridade até 2022

A RFEPCCT até 31 de dezembro de 2021 possuía 42 unidades, distribuídas por 661 localidades, espalhando-se por mais de 10% dos 5.570 municípios brasileiros (Brasil, 2019b; Ibge, 2019) e está presente em todos os 26 estados brasileiros, mais o Distrito Federal. Seu quadro de funcionários constitui-se em mais de 81 mil profissionais, compostos por mais de 46 mil professores e 42 mil técnicos (Brasil, 2019b; 2022b). Em 2018 havia aproximadamente 1 milhão de alunos matriculados e em 2021 já contava com mais de 1 milhão e meio de alunos matriculados (Brasil, 2022b).

O país possui um território segundo o Instituto Brasileiro de Geografia e Estatística (IBGE) de 8.516.000 km<sup>2</sup>. As 661 unidades da RFEPCCT estão distribuídas por esse território, havendo uma concentração maior na região litorânea. Essa distribuição acompanha a distribuição populacional do território brasileiro. Ela está presente em 100%

das 5 macrorregiões; aproximadamente em 97% das 137 mesorregiões; aproximadamente em 64% das 558 microrregiões (Brasil, 2016c; Ibge, 2022). A RFEPCT abrange aproximadamente 87% dos 118 territórios da cidadania (Brasil, 2016c). Por fim, abrange uma formação que vai do ensino infantil a pós-graduação de acordo com a plataforma – Nilo Peçanha – do governo federal brasileiro, a qual reuni os dados educacionais e administrativos da RFEPCT (Brasil, 2022b).

Conforme constatado a RFEPCT apresenta características diferentes de muitas instituições de ensino superior nas Américas, na Europa, na Ásia, na África e na Oceania (Brasil, 2016b). Essa história centenária reafirma a relevância e o escopo dessa rede singular de educação, pesquisa, extensão e interação com as demandas dos setores públicos e privados e governo, desde 2008, como uma das finalidades previstas legalmente.

Isoladamente, sua efetividade, no tocante ao ensino, fica evidente, ao se analisar isoladamente seus resultados no Programa Internacional de Avaliação de Estudantes (PISA). Resultados esses que, se apenas a RFEPCT fosse considerada, representando o Brasil, foram extraídas, entre as 10 IEPCT melhores avaliadas pelo critério estabelecido doravante denominada de (IEPCT-MACE), nos mesmos níveis de países como Japão, Canadá dentre outros e acima da média da OCDE (Corbucci, 2011; Gomes et al., 2019; Ramos, 2018).

Os propósitos e objetivos norteadores da RFEPCT estão em consonância com o modelo TH (Etzkowitz *et al.*, 2000), que apresenta como tese que as instituições de ensino superior podem desempenhar um papel importante no estímulo à inovação e ao desenvolvimento econômico regional.

É importante destacar que muito antes do surgimento dos conceitos da Quádrupla Hélice (Carayannis *et al.*, 2009), que agrega as demandas da “sociedade civil” e do “público baseado em mídia e cultura”; e da Quíntupla Helix (Bernardino *et al.*, 2020), que estende a “arquitetura da inovação ao meio ambiente” (sustentabilidade) e à “ecologia social”, a RFEPCT já apresentava esses fundamentos na lei que a instituiu em 2008, em seu artigo 6º, itens IV, VII e IX (finalidades).

Destaca-se, dentre as finalidades, a de consolidar e fortalecer os arranjos produtivos, sociais e culturais locais (Item IV), e de realizar e estimular, dentre outras ações, a produção cultural (Item VII), as quais passaram a ser objetos de estudos da Quádrupla Hélice em 2009. Já a finalidade de promover, desenvolver e transferir soluções tecnológicas sociais, especialmente direcionadas à preservação do meio ambiente e sua



sustentabilidade (Item IX), em 2010 passou a ser objeto de estudo da Quíntupla Helix. Isso posto, é possível identificar mais uma proposta vanguardista sobre a RFEPCT.

Nesta tese abordar-se-á os resultados das atividades relacionadas à produção intelectual, científica e tecnológica buscando enfatizar os resultados decorrentes de possíveis interações ICTs e SPP, cabendo a “hélice” governo (GOV) fomentar e garantir as condições para que essa sinergia possa efetivamente resultar no desenvolvimento local, regional e ou nacional desejado.

## Apêndice B - Quadros, Tabelas etc.

Quadro B. 1: Apresentação das principais plataformas públicas e privadas disponíveis até 2023 que buscam organizar os dados administrativos, financeiros, educacionais, de pesquisa e de extensão relacionados disponibilizados pelas ICTs.

<b>Id</b>	<b>Rótulo</b>	<b>Aplicação BI</b>	<b>Link TH</b>	<b>Descrição</b>
<b>P1</b>	Nilo Peçanha	Educacional, Administrativa e Financeira	ICTs	A aplicação BI tem como objetivo coletar, validar e disseminar as estatísticas da Rede Federal dos dados administrativos, financeiros e educacionais
<b>P2</b>	Painel Universidade 360	Educacional, Administrativa e Financeira	ICTs	A aplicação BI tem como objetivo disponibilizar dados integrados e indicadores acadêmicos, orçamentários e de gestão de pessoas das Universidades Federais
<b>P3</b>	Observatório Nacional de Ciência, Tecnologia e Inovação	Ecosistema nacional de CT&I	GOV e SPP	A aplicação BI tem como objetivo monitorar o estado da arte e as tendências relacionadas ao ambiente de CT&I, no Brasil e no mundo, para o Sistema Nacional de Ciência, Tecnologia e Inovação (SNCTI), por meio de KPIs e outros tipos de serviços de inteligência desenvolvidos pelo Centro de Gestão e Estudos Estratégicos, Organização Social do Ministério da Ciência, Tecnologia e Inovações (CGEE/MCTI)
<b>P4</b>	Portal Integra	Administrativa, Financeira e de P&D	ICTs	A aplicação BI desenvolvida pelo Instituto Federal do Rio Grande do Sul (IFRS), tem como objetivo reunir em um só lugar todas as habilidades e <i>expertise</i> dos funcionários; patentes ou tecnologia transferíveis; A infraestrutura para apoiar projetos de inovação; e informes sobre o desenvolvimento tecnológico das instituições de rede
<b>P5</b>	Stela Experta®	Administrativa, P&D e Grupos de Pesquisa	ICTs	Esta plataforma tem como objetivo oferecer às instituições de ensino superior brasileiras serviços estratégicos de apoio nas áreas de gestão do ensino, pesquisa, extensão e inovação, integrando automaticamente os dados dos plataforma LATTES e do diretório do grupo de pesquisa, bem como seus próprios dados a cada ICTs, a fim de apoiar a implementação de políticas de gestão

Nota: A aplicação Stela Experta® foi desenvolvida pelo Instituto Stela. O Instituto Stela é organização de P&D resultante do processo de *startup* proposta pelo grupo de pesquisa Stela da Universidade Federal de Santa Catarina. Observa-se que o grupo Stela desenvolveu a plataforma LATTES para o CNPq entre 1997 e 2004 (Marchezan *et al.*, 2012)

Quadro B. 2: Comparação das Principais Teorias que foram propostas para investigar os diversos tipos de interação entre as ICTs-SPP, com o objetivo de avaliar a natureza e o impacto dessas relações.

<b>Categorias</b>	<b>Sistemas Nacionais de Inovação</b>	<b>Teoria dos Sistemas de Inovação</b>	<b>Hélice Tríplice da Inovação</b>	<b>Teoria da Gestão do Conhecimento</b>	<b>Teoria da Inovação Aberta</b>
<b>Foco Principal</b>	Contexto nacional dos processos de inovação	Interações e redes em inovação	Interações entre Universidade, Indústria e Governo	Processos organizacionais de conhecimento	Fluxo de conhecimento através das fronteiras organizacionais
<b>Principal Proponente</b>	Christopher Freeman, Bengt-Åke Lundvall	Christopher Freeman, Bengt-Åke Lundvall	Henry Etzkowitz & Loet Leydesdorff	Ikujiro Nonaka, Hirotaka Takeuchi	Henry Chesbrough
<b>Artigo/Obra Seminal</b>	Freeman, C. (1987). Política Tecnológica e Desempenho Econômico: Lições do Japão.	Lundvall, B. Å. (Ed.). (1992). Sistemas Nacionais de Inovação: Rumo a uma Teoria de Inovação e Aprendizado Interativo.	Etzkowitz, H., & Leydesdorff, L. (2000). A dinâmica da inovação: de Sistemas Nacionais e “Modo 2” para uma Hélice Tríplice de relações universidade-indústria-governo.	Nonaka, I., & Takeuchi, H. (1995). A Empresa Criadora de Conhecimento.	Chesbrough, H. (2003). Inovação Aberta: O Novo Imperativo para Criar e Lucrar com Tecnologia.
<b>Papel da Colaboração Externa</b>	Importante	Central	Central	Incentivado	Central
<b>Tipos de Conhecimento</b>	Tácito e Explícito (implícito)	Tácito e Explícito (implícito)	Não especificamente definido	Tácito e Explícito	Tácito e Explícito (implícito)
<b>Papel das Instituições</b>	Central	Central	Importante	Normas Culturais e Organizacionais	Influente (ex., gestão de PI)
<b>Ênfase Tecnológica</b>	Frequentemente enfatizado	Infraestrutura (física e de conhecimento)	Não especificamente definido	Sistemas e ferramentas de GC	Enfatizado (ex., ferramentas de PI)
<b>Aspectos Culturais</b>	Cultura e história nacionais	Contexto histórico e cultural do sistema	Implicado (ex., cultura de colaboração)	Cultura de compartilhamento de conhecimento	Abertura para ideias externas

Continua...

Continua...

<b>Categorias</b>	<b>Sistemas Nacionais de Inovação</b>	<b>Teoria dos Sistemas de Inovação</b>	<b>Hélice Tríplice da Inovação</b>	<b>Teoria da Gestão do Conhecimento</b>	<b>Teoria da Inovação Aberta</b>
<b>Implicações de Política</b>	Central para moldar capacidades nacionais de inovação	Políticas moldando a capacidade de inovação do sistema	Influente na formação de colaborações	Políticas organizacionais para compartilhamento e retenção de conhecimento	Políticas de PI, políticas colaborativas
<b>Aprendizado e Adaptação</b>	Enfatizado	Central	Implicado através da colaboração	Central	Incentivado através da colaboração externa
<b>Mecanismos de Feedback</b>	Implicado na dinâmica do sistema	Central para a dinâmica do sistema	Implicado através das interações	Através dos processos de GC	Através de colaborações externas
<b>Top 10 Variáveis</b>	1. PIB 2. Investimento em P&D 3. Patentes 4. Publicações científicas 5. Colaborações internacionais 6. Nível educacional 7. Infraestrutura tecnológica 8. Políticas de inovação 9. Setores de alta tecnologia 10. Exportações de alta tecnologia	(Similar to NSI given the overlap)	1. Parcerias universidade-indústria 2. Spin-offs 3. Patentes conjuntas 4. Publicações conjuntas 5. Investimento em P&D 6. Políticas de inovação 7. Incubadoras 8. Transferência de tecnologia 9. Licenciamento 10. Startups	1. Fluxo de conhecimento 2. Práticas de compartilhamento 3. Sistemas de informação 4. Cultura organizacional 5. Estratégias de aprendizado 6. Retenção de conhecimento 7. Transferência de conhecimento 8. Capital intelectual 9. Inovação 10. Eficiência organizacional	1. Fluxo de ideias 2. Licenciamento 3. Patentes 4. Parcerias estratégicas 5. Modelos de negócios 6. P&D externo 7. Crowdsourcing 8. Co-criação 9. Ecossistemas de inovação 10. Startups

Nota: A seção "Top 10 Variáveis" fornece uma visão geral das variáveis comuns associadas a cada teoria. Essas variáveis podem variar de acordo com estudos ou contextos específicos.

Quadro B. 3: Avaliação original dos especialistas em relação ao MVP proposto.

Especialista	Ecossistema	Avaliação
1	ICTs/GOV	<p>A ferramenta PD&amp;S traz uma série de indicadores importantes sobre a performance dos institutos federais de ensino de modo a permitir a construção de diagnósticos, prognósticos e cenários diversos capazes de instruir os gestores das instituições sobre as melhores políticas a serem implementadas de modo a corrigir as políticas e práticas institucionais em andamento visando a melhoria do desempenho das instituições em seus variados aspectos. Tal ferramenta permite auscultar o status de cada instituto no tocante a seus feitos seja na pesquisa, seja em P&amp;D, em publicação, orientação de alunos e, ainda, sobre o montante de recursos mobilizados para cada ação, bem como de seus correlatos resultados/produtos finais. A ferramenta é passível de ser utilizada e customizada para aferir e monitorar outras ações de outras instituições para além dos institutos federais e com aderência suficiente para, por exemplo, captar os indicadores relacionados à pós-graduação. Atualmente, como diretor científico e de inovação da Fundação de Amparo à Pesquisa do Estado de Alagoas (FAPEAL), vejo a efetividade desta ferramenta quando a comparo com o sistema adotado pela FAPEAL na gestão de seus projetos científicos (SparkX) do Instituto Stela. E, frise-se, há um diferencial: esta ferramenta (DP&amp;S) rastreia dados e indicadores que não existe em nossa ferramenta, sobre, por exemplo, projetos de desenvolvimento, treinamento, consultoria, extensão acadêmica e marcas. De natureza quantitativa, esta ferramenta permite a estruturação da análise qualitativa referente às categorias principais. A solução da ferramenta PD&amp;S é relevante e, pode ser ainda mais efetiva se ajustada às necessidades de agências de fomento à pesquisa, como a FAPEAL.</p>

Continuação...

Continuação...

Especialista	Ecossistema	Avaliação
2	ICTs/GOV	<p>Com o objetivo de concentrar os dados sobre a Rede Federal de Educação Profissional Científica e Tecnológica (RFEPCT) gerando informações que possam contribuir com os gestores em suas tomadas de decisões, como também dar suporte a servidores, pesquisadores, discentes e a comunidade em geral em seus mais diversos tipos de pesquisas, criamos um painel visual (<i>dashboard</i>) que traz a análise de dados de ICTs pertencentes à RFEPCT brasileira permitindo análises para além dos modelos tradicionais existentes educacionais e de gestão. O <i>dashboard</i> criado é um sistema de <i>Business Intelligence</i> (BI), o qual utilizou, enquanto objeto de aplicação, os dados de pesquisa, desenvolvimento e prestação de serviços científicos e/ou tecnológicos coletados de várias ICTs, num determinado horizonte de tempo, que pode ser expandido e atualizado constantemente. O <i>dashboard</i> concentra em uma única plataforma os dados coletados na Lei orçamentária anual, Portal da Transparência, Portal do Censo Superior da Educação, Diretório dos Grupos de Pesquisa do CNPQ, Plataforma Lattes, Instituto Nacional de Propriedade Industrial, além de Relatórios de gestão das ICTs da RFEPCT.</p> <p>Questões a serem avaliadas: Relevância da solução proposta; comparada a outras soluções semelhantes, quais suas vantagens e desvantagens; Avaliação da ferramenta e seu potencial de adoção por um gestor; Sugestões/críticas. Link do objeto: <a href="https://prequest.websiteseuro.com/bi">https://prequest.websiteseuro.com/bi</a></p> <p>Das considerações:</p> <p>A ferramenta apresenta as seguintes características positivas: Design amigável; Boa navegabilidade/iteração com alguns dos gráficos apresentados; Página inicial adequada e sugestiva para que o usuário entenda o conteúdo a ser encontrado nas próximas visualizações; Versões em português e inglês; Conjunto de informações relevantes para o ecossistema de CT&amp;I; As visões disponíveis trazem um conjunto de KPI's ou métricas bastante relevantes para as análises por parte de uma gestão: Consultoria; Extensão Acadêmica; Outras atividades; Marcas; Desenho Industrial; Patente Publicada; Programa de Computador.</p> <p>Sugestões de melhoria:</p> <p>No quesito de usabilidade: Os botões laterais e superiores (exceto ícone de volta à página principal), não apresentam características intuitivas/explicativas (apresentam apenas um <i>label</i> com texto “Clique aqui para seguir link”; A mensagem fixa “Aguarde o carregamento no canto superior esquerdo após a aplicação de um filtro”, não é clara em seu propósito; Há menções diferentes para o combo de seleção: Filtro e Variável; Verificar a possibilidade da inserção de um manual online de uso para a ferramenta; Verificar a possibilidade de incluir um guia de referência metodológica para as visões disponíveis</p> <p>Resumo da avaliação:</p>

	<p>A ferramenta PD&amp;S apresenta um bom conjunto de informações, que auxiliam na avaliação e no acompanhamento das instituições da rede Federal de EPCT. Esse conjunto de informações pode fornecer subsídios para tomada de decisão, formulação/avaliação de políticas implementada e, por conseguinte auxiliar no aprimoramento da gestão e melhoria da qualidade das ações desenvolvidas por essas instituições.</p> <p>A ferramenta tem um propósito inicial focado nas ações de CT&amp;I, de Pesquisa e Extensão, mas pode também contemplar ações voltadas à educação, o que auxiliaria a consolidação desse tripé.</p> <p>Além disso, a customização da ferramenta para contemplar outras instituições que compõe o ecossistema de CT&amp;I do Brasil, poderia tornar a ferramenta ainda mais relevante.</p> <p>Avaliar as características de demandas das Fundações de Amparo à Pesquisa (FAPs) dos estados, pode trazer também novas possibilidades para a ferramenta.</p> <p>Como gestor de inovação da SECTI-BA, temos iniciativa similar <a href="https://infovis.sei.ba.gov.br/cti/">https://infovis.sei.ba.gov.br/cti/</a>, inserido dentro de um portal mais amplo e um observatório de CT&amp;I em desenvolvimento (que deve ser incorporado ao INFOVIS).</p> <p>A FAPESB utiliza ferramenta que está em processo de substituição por uma solução que incorpora elementos de Inteligência Artificial, para cumprir e aprimorar a operação e a gestão atual.</p> <p>Finalmente, avalio que a ferramenta PD&amp;S apresenta uma relevância considerável para o propósito apresentado e já pela versão disponível e que possui um potencial de melhoria e ampliação de futuros clientes da solução.</p>
--	--

Continuação...

Continuação...

Especialista	Ecosistema	Avaliação
3	ICTs	<p>Primeiramente eu agradeço a oportunidade de testar a ferramenta e com isso opinar, contribuindo para o seu desenvolvimento.</p> <p>A ferramenta será de grande utilidade para a Rede federal de Ensino Profissional devido às questões da área de pesquisa e inovação.</p> <p>Um portal com todas as informações referentes aos potenciais em números da rede federal contribui não só para pesquisa, mas também para divulgação das ações da rede na área de inovação.</p> <p>Como gestor, eu também cito a relevância do sistema para a tomada de decisão e ampliação das áreas de pesquisa de cada instituição, além de promover parcerias.</p> <p>Para finalizar, eu gostaria de incluir alguns pontos para a pesquisa:</p> <ol style="list-style-type: none"><li>1. Seria importante no próprio portal ter a informação de referência das fontes e dos bancos de dados;</li><li>2. Poderia implantar um modelo de maturidade de inovação baseado em alguns dados para a avaliação de uma determinada instituição;</li><li>3. Uma questão que temos dificuldade na instituição é filtrar os pesquisadores da Rede pela sua área ou temática de pesquisa;</li><li>4. Caracterizar a fonte de recurso seria importante também;</li><li>5. Lei de informática e outras leis poderiam ser incluídas?</li></ol>

Continuação...



Continuação...

Especialista	Ecossistema	Avaliação
4	SPP	<p>Avalie a relevância da solução proposta e se ela atende o objeto proposto ou não;  Resposta: Diante da análise do MVP encaminhado venho afirmar que a solução proposta possibilita não só atender a análise objetiva e confiável dos canais de PD&amp;S em ecossistema das ICTs, utilizando dados obtidos a partir de múltiplas fontes, a fim de fundamentar a tomada de decisão, análise, gestão e monitoramento das demandas do setor produtivo e do governo, como também avaliar a potencialidade do capital empreendedor e seus respectivos estoques de capacidade, pois o produto apresentado demonstra seu potencial ao expor os produtos, processos e serviços vinculados a cadeia denominada de PD&amp;S de forma simples, objetiva e clara.</p> <p>Avalie a ferramenta e seu potencial de adoção por um gestor;  Resposta: A ferramenta permite que um gestor possa navegar pela ferramenta e desenvolver análises diretas a partir dos gráficos e filtros disponíveis, facilitando que um gestor do setor produtivo possa identificar oportunidades dentro da academia para potencializar seu leque de negócios, identificando produtos, processos e ou serviços, como por exemplo consultorias sob demanda para resolver problemas de seus parceiros comerciais ou da própria instituição. Outro ponto a ser destacado é a geração de projetos em âmbito nacional no desenvolvimento do empreendedorismo (foco SEBRAE). O MVP permite que o gestor após realizar o filtro desejado exporte os dados para outras ferramentas de estatística para desenvolver análises mais complexas, porém para acesso a essa funcionalidade foi necessário solicitar acesso offline ao MVP, dada a limitação gratuita que foi disponibilizado. Outro ponto detectado foi a facilidade de integrá-la a outras plataformas.</p> <p>Sua viabilidade comercial;  Resposta: Sim, porém em se tratando de um MVP é necessário investimento para torná-la comercializável, agregando outras funcionalidades como destaque de potenciais áreas de desenvolvimento e pesquisa de acordo com a demanda de mercado.</p> <p>Sugira/critique a ferramenta com base nas métricas disponíveis lembrando que são sugestões iniciais e que podem ser adequadas às necessidades dos gestores;  Resposta: Sim, porém em se tratando de um MVP é necessário investimento para torná-la comercializável, para torná-lo mais simplificado a nível de usuário com pouca compreensão da área de dados. Outro ponto é a atualização online dos dados de forma dinâmica.</p> <p>Aponte sugestões de melhoria e inovação para a ferramenta;  Resposta: Sim, porém em se tratando de um MVP é necessário investimento para torná-la comercializável, agregando a funcionalidade de integração de uma API de serviços de entrega de dados. Inclusão de um painel para gerenciamento de usuários a nível de desenvolvimento e apresentação dos dados.</p>

		<p>Faça uma comparação desta ferramenta com outras soluções semelhantes existentes, caso conheça, apontando suas vantagens e desvantagens;</p> <p>Resposta: Sim, o Stela Experta. Ao comparar a solução apresentada com a única ferramenta similar de mercado (Stela Experta), permite comparações e interações que a ferramenta comercial não disponibilizar dentro da proposta do MVP apresentado, sem falar que o acesso ao MVP é público e gratuito.</p>
5	ICTs	<p>Seguem minhas avaliações:</p> <p>A solução apresentada tem grande relevância, particularmente por conta da escassez de informação acerca dos diversos dados relevantes para avaliar e agir por sobre as políticas públicas que mantêm e aperfeiçoam as instituições de educação, ciência, tecnologia e inovação. Neste aspecto, a BI proposta vai na direção de prover um painel bastante rico, embora ainda necessite de ajustes na interface e melhor definição de indicadores.</p> <p>Gestores de políticas públicas em órgãos governamentais necessitam de rápido e confiável acesso a informações acerca de indicadores relevantes, tanto para identificar padrões de comportamento que permita a proposição e a modelagem de políticas novas quanto para a análise mais assertiva das políticas atuais. Da mesma forma, gestores institucionais necessitam interpretar o atual estágio de suas realidades locais e as capacidades e deficiências perante as demandas do mercado e do governo. A ferramenta proposta é um passo interessante nessa direção.</p> <p>A viabilidade comercial de uma ferramenta depende essencialmente da relevância do problema que ataca, da qualidade com que o faz e da exclusividade da forma com que a executa, de forma a manter uma significativa vantagem com relação a potenciais competidores. A ferramenta apresentada, neste último aspecto, não parece apresentar barreiras à concorrência, mas poderia ocupar nichos enquanto consolida vantagens maiores.</p> <p>Há métricas e indicadores que necessitam de melhor clareza, como, por exemplo, "pesquisadores por ano", o que não parece de forma alguma, pelo comportamento altamente variável dos valores, o que o nome sugere (o número de pesquisadores pertencentes às instituições em cada ano). Certamente o número de pesquisadores nos IFs não passou de mais de 10.000 em 2008 para menos de mil em 2009, etc. Fragilidades como esta ou diferentes aparecem em outros itens, como por exemplo o caso do "investimento federal em P&amp;D (R\$) por ano". Claramente os valores mostrados não guardam relação com o que o nome significa. Isso necessita ser revisto em cada caso para que haja clareza, ou ninguém utilizará a ferramenta para seu objetivo.</p> <p>A interface precisa ser melhor organizada e deve-se diminuir o esforço visual para a identificação dos indicadores e valores. Isso, além de uma mais adequada definição das informações mostradas, de modo a trazer confiança, são as principais alterações necessárias. Depois deve-se trabalhar com a inclusão de uma maior variedade de ICTs.</p> <p>Desconheço ferramentas exatamente semelhantes.</p>

Continuação...

Continuação...

Especialista	Ecossistema	Avaliação
6	SPP	<p>Avalie a relevância da solução proposta e se ela atende o objeto proposto ou não; Atende certamente, será de grande utilidade pública.</p> <p>Avalie a ferramenta e seu potencial de adoção por um gestor; Total, pode trazer fomentos a muitas políticas públicas e potenciais melhorias nos sistemas de ensino.</p> <p>Sua viabilidade comercial; Há viabilidade de comercialização com certeza, de baixo custo.</p> <p>Sugira/critique a ferramenta com base nas métricas disponíveis lembrando que são sugestões iniciais e que podem ser adequadas às necessidades dos gestores;</p> <p>Faltou um manual de utilização do dashboard, essa parte parece não ter preenchimento.</p> <p>Aponte sugestões de melhoria e inovação para a ferramenta;</p> <p>Criar mais abas laterais de navegação, ficou muito extenso o dashboard, ordenar os gráficos de barras que não tiverem uma ordenação natural, como ano, por exemplo. Para melhor visualização dos dados.</p> <p>Faça uma comparação desta ferramenta com outras soluções semelhantes existentes, caso conheça, apontando suas vantagens e desvantagens;</p> <p>Não conheço, acho que é uma excelente inovação.</p>

Continuação...

Continuação...

Especialista	Ecossistema	Avaliação
8	ICTs	<p>1. A concepção da ferramenta em si é de grande relevância, uma vez que permitirá aos gestores, ou seja, reitores e pró-reitores, ter uma visão macro da sua instituição na sua vertente maior que é PD&amp;I. Neste ponto, senti falta na ferramenta de um ambiente que permita ver especificamente apenas uma instituição de interesse bem como ver gráficos comparativos entre instituições. A ideia seria permitir responder perguntas tais como: Como está a instituição que eu dirijo com relação à média brasileira? Ou como comparar em algum item a minha instituição com outra do mesmo porte? Como saber dos investimentos aqueles que foram estaduais (via FAPs, por exemplo) e os Federais?</p> <p>Acredito que esse perspectiva a ferramenta deveria atender.</p> <p>2. Outro ponto importante é entender o conceito de pesquisa acadêmica (e estendo aos outros itens de avaliação), pois questiono isso uma vez que na ferramenta, quando se coloca “orientação de pesquisa”, e em “gráficos por título ou descrição” aparece “Estágio supervisionado”, “Estágio curricular obrigatório” etc que não é orientação de pesquisa, fica estranho (pode ser minha internet ruim neste momento...) mas fica o alerta para uma possível revisão e melhor entendimento do que se espera. Inclusive, em orientação a pesquisa, como em pesquisa, um ponto que deve ser associado é a questão de publicação no período da pesquisa associada a instituição. Nesta questão pesquisa também neste mesmo item “gráficos por título ou descrição” vale verificar o tipo de trabalho (avalei apenas pelo título) que aparece porque lá tem inclusive “Criação de catálogo de moda” e também aparecem os programas que não necessariamente são de pesquisa (mais uma vez, pode ser problema meu de conexão) mas sugiro rever a conexão entre este item e o que se espera quando se faz a seleção da variável a analisar.</p> <p>Na variável “Patente publicada” em “tempo médio para publicação” está saindo 2,05 o que conhecendo o INPI sabemos que a média é de 10 anos. Não sei como esse parâmetro foi gerado, mas recomendo rever e consequentemente rever os outros tipos proteções.</p> <p>Em resumo, com relação aos itens solicitados:</p> <ul style="list-style-type: none"><li>- Avalie a relevância da solução proposta e se ela atende o objeto proposto ou não;</li></ul> <p>A solução é excelente e a meu ver ela atenderia sim ao objetivo proposto.</p> <ul style="list-style-type: none"><li>- Avalie a ferramenta e seu potencial de adoção por um gestor: A ferramenta tem um grande potencial para ser adotada por um gestor, inclusive presumo que possa ser utilizada no contexto de ICTs com os devidos ajustes e instanciamentos. Requer apenas de ajustes na concepção para ter gráficos comparativos e correlações entre as variáveis iniciais e os resultados apresentados</li></ul>

		<p>- Sua viabilidade comercial: É viável para uma transferência de tecnologia, através de um licenciamento, uma vez que cada instituição que deseje utilizada precisará ter um contrato de adequação das funcionalidades da ferramenta a realidade da instituição, do estado, etc.</p> <p>- Sugira/critique a ferramenta com base nas métricas disponíveis lembrando que são sugestões iniciais e que podem ser adequadas às necessidades dos gestores: já feito anteriormente/inicialmente</p> <p>Aponte sugestões de melhoria e inovação para a ferramenta: já feito anteriormente/inicialmente</p> <p>- Quanto ao último item, a comparação, não conheço soluções semelhantes.</p>
--	--	---

Continuação...

Continuação...

Especialista	Ecossistema	Avaliação
9	ICTs	<p>2- O produto é relevante para a gestão e acompanhamento de ações/processos de pesquisa e atende os objetivos a que se propõe;</p> <p>3- Do ponto de vista da gestão, o produto apresentado promove um excelente mapeamento de informações importantes que pode facilitar a tomada de decisões e direcionamento de ações diretas de aprimoramento dos processos de pesquisa, desenvolvimento e inovação das instituições;</p> <p>4- Apresenta potencial para viabilidade comercial;</p> <p>5- A ferramenta contempla pontos essenciais para a gestão, análise e acompanhamento de processos relativos à pesquisa em ICTs;</p> <p>6- No momento, sem sugestões. Creio que, com o tempo, novas necessidades surgirão e deverão ser inseridas. Neste sentido, quero somente lembrar que a ferramenta deve estar adequada para a inserção de novas variáveis e parâmetros que se fizerem necessários para atender a realidades que se impuserem com o passar do tempo;</p> <p>7- O produto tecnológico encontra-se bem amplo, apresentando informações que outras soluções no mesmo sentido não apresentam, a exemplo do investimento destinado a cada instituição, que se constitui uma informação fundamental para a gestão e tomada de decisão relacionadas ações/ processos que envolvem a Pesquisa e a Inovação.</p>

Quadro B. 4: Síntese Evolutiva do Modelo Tripla Hélice 1958 – 2022

<b>Topologia Triple Helix</b>	<b>I</b>	<b>II</b>	<b>III</b>	<b>IV</b>	<b>V</b>
<b>Ano</b>	1958	1968	2000	2009	2010
<b>Artigo Semente</b>	(Mills, 1958)	(Sabato <i>et al.</i> , 1968)	(Etzkowitz <i>et al.</i> , 2000)	(Carayannis <i>et al.</i> , 2009)	(Carayannis <i>et al.</i> , 2010)
<b>Modo Conhecimento</b>	-	Modo 2	Modo 3	Modo 3	Modo 3
<b>Abordagem</b>	O Estado-nação abrange academia e indústria e dirige as relações entre eles (Etzkowitz <i>et al.</i> , 2000, p. 111-112)	Consiste em esferas institucionais separadas com fronteiras fortes dividindo-se eles e relações altamente circunscritas entre as esferas. (Etzkowitz <i>et al.</i> , 2000, p. 111-112)	Consiste na superposição de esferas institucionais, cada uma assumindo o papel da outra e com organizações híbridas emergentes nas interfaces está gerando uma infraestrutura. (Etzkowitz <i>et al.</i> , 2000, p. 111-112)	Destaca um Ecossistema de Inovação que estimula a co-evolução entre diferentes fontes de conhecimento e inovação e equilibra formas não-lineares de inovação em sistemas multinível. (Carayannis <i>et al.</i> , 2009, p. 227)	Enfatiza a transição socioecológica necessária da sociedade e da economia no século XXI
<b>Modelo</b>	Falho, onde o papel do Estado-nação define e controla o processo de inovação (Etzkowitz <i>et al.</i> , 2000, p. 111-112)	Utiliza uma política de <i>laissez-faire</i> , buscando reduzir o papel do Estado-nação defendido pelo modelo TH I (Etzkowitz <i>et al.</i> , 2000, p. 111-112)	Arranjos UIG são incentivados, mas não regulados pelo governo. A regulamentação pode ser através de regras, assistência financeira, lei Bayh-Dole nos EUA, ou novos atores como fundações de inovação da Suécia. (Etzkowitz <i>et al.</i> , 2000, p. 112)	Somasse aos arranjos UIG, a cultura. A cultura influencia todos os SNI. A própria "cultura da inovação" é fundamental para promover uma economia avançada baseada no conhecimento. (Carayannis <i>et al.</i> , 2009, p. 227)	Somasse aos arranjos UIGC o meio ambiente. São considerados o meio ambiente social e econômico para a produção do conhecimento e os sistemas de inovação

Continua...

Continua...

Topologia Triple Helix	I	II	III	IV	V
<b>Objetivo</b>	A centralização burocrática do Estado-nação é o principal impulso organizador da inovação.	Transformar a convergência dos arranjos (University-Industry-Government) em centros capazes de gerar, incorporar e transformar demandas em um produto final que é a inovação científico-tecnológica. (Sabato <i>et al.</i> , 1968, p. 6)	Desenvolver um ambiente inovador composto por empresas <i>spin-off</i> universitárias, iniciativas trilaterais para o desenvolvimento econômico baseado no conhecimento e alianças estratégicas entre empresas grandes e pequenas, operando em diferentes áreas e com diferentes níveis de tecnologia, laboratórios governamentais, e grupos de pesquisa acadêmica. (Etzkowitz <i>et al.</i> , 2000, p. 112)	Integrar a perspectiva do público baseado na mídia e na cultura resultando num ecossistema emergente de conhecimento e inovação fractal, bem configurado para a economia do conhecimento e a sociedade. (Carayannis <i>et al.</i> , 2009, p. 201)	Integrar os ambientes naturais da sociedade e da economia como impulsionadores da produção e inovação do conhecimento, definindo assim oportunidades para a economia do conhecimento

Nota. Os tópicos de abordagem e modelo foram referenciados com base nos relatos de (Etzkowitz *et al.*, 2000, p. 111-112; Galvão *et al.*, 2019) e (Carayannis *et al.*, 2009; Carayannis *et al.*, 2010; Carayannis *et al.*, 2022). A expressão francesa *laissez-faire* significa “deixar fazer”, sendo considerada uma expressão simbólica da economia liberal defendida pelo capitalismo.

Quadro B. 5: Principais variáveis utilizadas para analisar os Canais de PD&S em ICTs no contexto da Teoria TH

<b>Variáveis relacionadas ao canal institucional</b>			
<b>ID</b>	<b>Canais</b>	<b>Variáveis</b>	<b>Aporte teórico</b>
1	A	desmotivação dos funcionários	(Niwa, 2014)
2	A	excesso de burocracia	(Niwa, 2014)
3	A	formações dos pesquisadores	(Miano; Do Couto; <i>et al.</i> , 2017)
4	A	lentidão nos procedimentos	(Niwa, 2014)
5	A	modelo administrativo adotado	(Niwa, 2014)
6	A	número alunos de matriculados	(Miranda <i>et al.</i> , 2018)
7	A	número de docentes	(Miranda <i>et al.</i> , 2018)
8	F	bolsas de iniciação científica	(Santos <i>et al.</i> , 2015)
9	F	financiamentos/investimento	(Berbegal-Mirabent <i>et al.</i> , 2019) (Cesaroni <i>et al.</i> , 2016) (Menezes <i>et al.</i> , 2019) (Miranda <i>et al.</i> , 2018) (Rapini <i>et al.</i> , 2019) (Šereš <i>et al.</i> , 2019)
<b>Variáveis relacionadas aos canais de Desenvolvimento tecnológico</b>			
<b>ID</b>	<b>Canais</b>	<b>Variáveis</b>	<b>Aporte teórico</b>
10	D	patentes	(Basso <i>et al.</i> , 2021); (Berbegal-Mirabent <i>et al.</i> , 2019); (Cesaroni <i>et al.</i> , 2016) (Chaves <i>et al.</i> , 2017) (D'este <i>et al.</i> , 2007) (Macedo, 2014) (Menezes <i>et al.</i> , 2019) (Perucchi <i>et al.</i> , 2015) (Perucchi <i>et al.</i> , 2017)
11	D	patentes SPP	(D'este <i>et al.</i> , 2007).
12	D	produção tecnológica	(Chaves <i>et al.</i> , 2017); (Garcia, 2011); (Perucchi <i>et al.</i> , 2017);

Continua...



Continua...

<b>Variáveis relacionadas ao canal de Pesquisa</b>			
<b>ID</b>	<b>Canais</b>	<b>Variáveis</b>	<b>Aporte teórico</b>
1	A	desmotivação dos funcionários	(Niwa, 2014)
13	P	apoio institucional para atividades de pesquisa	(Perucchi <i>et al.</i> , 2015)
14	P	apresentações de trabalho	(Santos <i>et al.</i> , 2015)
15	P	áreas de concentração dos GP	(Miano; Do Couto; <i>et al.</i> , 2017)
16	P	artigos completos publicados em periódicos	(Berbegal-Mirabent <i>et al.</i> , 2019) (Perucchi <i>et al.</i> , 2017) (Ribeiro <i>et al.</i> , 2020) (Santos <i>et al.</i> , 2015) (Šereš <i>et al.</i> , 2019)
17	P	características dos GP	(Macedo, 2014) (Miano <i>et al.</i> , 2020)
18	P	consolidação do GP	(Macedo, 2014)
19	P	número de pesquisadores	(Berbegal-Mirabent <i>et al.</i> , 2019) (D'este <i>et al.</i> , 2007) (Jovanović <i>et al.</i> , 2022) (Šereš <i>et al.</i> , 2019)
20	P	orientação de pesquisa	(Berbegal-Mirabent <i>et al.</i> , 2019) (Sierra <i>et al.</i> , 2018)
21	P	pesquisa	(Menezes <i>et al.</i> , 2019) (Perucchi <i>et al.</i> , 2015)
22	P	produção científica	(Chaves <i>et al.</i> , 2017), (Fan <i>et al.</i> , 2015), (Garcia, 2011), (Menezes <i>et al.</i> , 2019), (Miano <i>et al.</i> , 2020), (Rapini <i>et al.</i> , 2019), (Ribeiro <i>et al.</i> , 2020), (Bento, 2015), (Santos <i>et al.</i> , 2015)

Continua...

Continua...

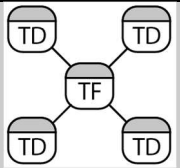
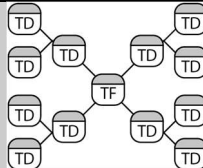
<b>Variáveis relacionadas aos canais de Desenvolvimento tecnológico</b>			
<b>ID</b>	<b>Canais</b>	<b>Variáveis</b>	<b>Aporte teórico</b>
23	P	projetos de pesquisa	(Berbegal-Mirabent <i>et al.</i> , 2019) (D'este <i>et al.</i> , 2007) (Šereš <i>et al.</i> , 2019) (Sierra <i>et al.</i> , 2018)
24	P	quantidade de Grupo de pesquisa	(Silva, 2015) (Miano; Do Couto; <i>et al.</i> , 2017)
25	P	resumos expandidos publicados em anais de congressos	(Perucchi <i>et al.</i> , 2017) (Santos <i>et al.</i> , 2015)
26	P	resumos publicados em anais de congressos	(Perucchi <i>et al.</i> , 2017), (Santos <i>et al.</i> , 2015)
27	P	trabalhos completos publicados em anais de congressos	(Berbegal-Mirabent <i>et al.</i> , 2019) (Perucchi <i>et al.</i> , 2017) (Santos <i>et al.</i> , 2015) (Šereš <i>et al.</i> , 2019)
28	P&D	apoio institucional para atividades de PD&I	(Silva, 2015)
29	P&D	inovação	(Perucchi <i>et al.</i> , 2015)
30	P&D	fatores propulsores do PD&I	(Silva, 2015)
31	P&D	projetos de P&D	(Berbegal-Mirabent <i>et al.</i> , 2019)
32	P&D	projetos de P&D financiados pelo SPP	(Cesaroni <i>et al.</i> , 2016)
33	PD&S	gestão do conhecimento	(Lino, 2013)
34	PD&S	interação TH	(Macedo, 2014), (Niwa, 2014)
35	PD&S	produção compartilhada do conhecimento	(Macedo, 2014) (Santos <i>et al.</i> , 2015)
36	PD&S	<i>spin-offs</i>	(Galati <i>et al.</i> , 2020)
37	PD&S	transferência de tecnologia	(Rapini <i>et al.</i> , 2019)
38	S	apoio institucional para atividades de extensão tecnológica	(Silva, 2015)

Continua...

Continua...

ID	Canais	Variáveis	Aporte teórico
39	S	consultorias	(D'este <i>et al.</i> , 2007) (Perucchi <i>et al.</i> , 2017) (Šereš <i>et al.</i> , 2019)
40	S	consultorias SPP	(Berbegal-Mirabent <i>et al.</i> , 2019) (D'este <i>et al.</i> , 2007) (Nsanzumuhire <i>et al.</i> , 2020)
41	S	licenças de tecnologia	(Tseng <i>et al.</i> , 2020)
42	S	prestação de serviços	(Lino, 2013) (De Negri, F. O. <i>et al.</i> , 2016) (Nascimento, 2021)
43	S	treinamentos	(D'este <i>et al.</i> , 2007)

Quadro B. 6: Principais diferenças entre esquemas estrela e floco de neve

Atributos	Estrela	Floco de neve
<b>Visual</b>		
<b>Modelo</b>	Simple	Complexo
<b>Consulta</b>	Simple	Complexa
<b>Performance</b>	Rápida	Lenta
<b>Manutenção</b>	Fácil	Complexa
<b>Integridade</b>	Risco de registro duplicados	Não risco de registro duplicados
<b>Nível</b>	Um nível de tabelas dimensão	Multinível de tabelas dimensão
<b>Relacionamento</b>	Chave primária da tabela dimensão relacionada a uma chave estrangeira na tabela fato	Chave primária da tabela dimensão relacionada a uma chave estrangeira na tabela fato e entre tabelas dimensão

Nota: Tabela fato (TF); Tabela dimensão (TD). Este quadro foi adaptado de Standardization *et al.* (2011) e Balakrishnan (2022).

Quadro B. 7: Resultado da revisão bibliométrica utilizado para elaboração do histórico da RFEPECT

<b>Termos de busca utilizados foram por ordem de relevância</b>	
Palavra-chave: i) "aprendizes artifices"; ii) "collegio pedro" ou "colegio pedro"; iii) "curso tecnico" ou "cursos tecnicos"; iv) "escola tecnica", "escolas tecnicas", "escolas industriais" ou "escola industrial"; v) "centro federal" ou "centros federais"; vi) "instituto federal", "institutos federais", "federal institutes" ou "federal institutions"; vii) "rede federal"; viii) "nilo pecanha" ou "nilo procopio pecanha"; ix) "educacao profissional", "educacion profesional", "education technology", "educacao industrial", "professional education" ou "professional education"; x) "universidade tecnologica"; xi) autonomia, contextualizacao, lei, decreto, norma ou historia.	
<b>Informações principais sobre os dados</b>	
<b>Período:</b>	1837:2022
<b>Fontes (Documentos governamentais, leis, decretos, livros, biblioteca de teses e dissertações da Capes etc.):</b>	24
<b>Palavras-chave dos autores:</b>	20
<b>Artigo:</b>	6
<b>Leis ou regulamentações e Documentos governamentais:</b>	18

Quadro B. 8: Resultado da revisão bibliométrica utilizado para elaboração do histórico da evolução dos canais de PD&S do ecossistema da RFEPECT

<b>Termos de busca utilizados foram por ordem de relevância</b>	
Palavra-chave: i) "colegio pedro", "pedro school" or "pedro college"; ii) "centro federal", "centros federais", "federal centers", "federal center"; iii) "instituto federal", "institutos federais", "federal institutes" ou "federal institutions"; iv) "rede federal" ou "federal network"; v) "universidade tecnologica" ou "technological university"; vi) pesquisa, "gestao do conhecimento", "propriedade intelectual", "transferencia de tecnologia", "producao cientifica", "indicadores de producao", "desenvolvimento tecnologico", inovacao, "incubadoras tecnologicas", servico ou serviços.	
<b>Informações principais sobre os dados</b>	
<b>Período:</b>	2008:2022
<b>Fontes (Jornais, Livros etc.):</b>	19
<b>Palavras-chave dos autores:</b>	37
<b>Artigo:</b>	19

Quadro B. 9: Resultado da revisão bibliométrica utilizado para contextualizar alguns fatores nacionais de CT&I e PD&I em que o ecossistema das ICTs se desenvolveu nas duas últimas décadas

<b>Termos de busca utilizados foram por ordem de relevância</b>	
Palavra-chave: i) <i>"triple helix" ou "tripla helice"</i> ; ii) <i>"economia do conhecimento" ou "knowledge economy"</i> ; iii) <i>"pesquisa e desenvolvimento" ou "research and development"</i> ; vi) <i>"desenvolvimento científico" ou "scientific development"</i> ; v) <i>"inovacao" ou "innovation"</i> ; vi) <i>"tecnologia" ou "technology"</i> ; vii) <i>"pesquisa" ou "research"</i> ; viii) <i>"ciencia" ou "science"</i> ; xi) <i>"marco legal" ou "legal frame"</i> ; x) <i>"productivity" ou "produtividade"</i> ; xi) <i>"metric" ou "metrica"</i> ; xii) <i>"science policy" ou "politica cientifica"</i> ; xiii) <i>"publicas publicas" ou "public policy"</i> .	
<b>Informações principais sobre os dados:</b>	
<b>Período:</b>	1995:2022
<b>Fontes (Jornais, Livros etc.):</b>	10
<b>Palavras-chave dos autores:</b>	16
<b>Artigos:</b>	6
<b>Leis ou regulamentações e Documentos governamentais:</b>	2

Quadro B. 10: Resultado da revisão bibliométrica utilizado para contextualizar a teoria da TH da inovação utilizada na investigação dos fenômenos dos canais de PD&S desenvolvidos no ecossistema das ICTs

<b>Termos de busca utilizados foram por ordem de relevância</b>	
Palavra-chave: i) <i>"analytical framework" ou "quadro analitico"</i> ; ii) <i>"research and development" ou "pesquisa e desenvolvimento"</i> ; iii) <i>"mode 2" ou "modo 2"</i> ; iv) <i>"university-industry" ou "universidade-industria" ou "universidade-empresa"</i> ; v) <i>"triple helix" ou "tripla helice"</i> ; vi) <i>"national systems innovation" ou "sistema nacional de inovação"</i> , vii) <i>"public services" ou "serviços publicos"</i> , e viii) <i>"systematic review" ou "revisão sistematizada"</i>	
<b>Informações principais sobre os dados:</b>	
<b>Período:</b>	1958:2022
<b>Fontes (Jornais, Livros etc.):</b>	20
<b>Palavras-chave dos autores:</b>	13
<b>Artigos:</b>	16
<b>Leis ou regulamentações e Documentos governamentais:</b>	3
<b>Livros:</b>	1

Quadro B. 11: Resultado da revisão bibliométrica utilizado para contextualizar a teoria da tripla hélice da inovação utilizada na investigação dos fenômenos dos canais de PD&S desenvolvidos no ecossistema das ICTs

<b>Termos de busca utilizados foram por ordem de relevância</b>	
Palavra-chave: i) <i>"knowledge brokerage"</i> ; ii) <i>"big data education"</i> ; iii) <i>"dashboard"</i> ; iv) <i>"ux design" ou "user experience"</i> ; v) <i>"business intelligence"</i> ; vi) <i>"acm proceedings"</i> ; vii) <i>"activity diagram"</i> ; viii) <i>"Data Warehouse"</i> ; ix) <i>"performance measurement"</i> ; x) <i>"Teoria da Decisão Racional" ou "Rational Decision Theory"</i> ; xi) <i>"Teoria da Gestão do Conhecimento" ou "Knowledge Management Theory"</i> ; xii) <i>"Teoria da Análise de Negócios" ou "Business Analysis Theory"</i> ; xiii) <i>"Teoria da Inteligência Competitiva" ou "Competitive Intelligence Theory"</i> ; xiv) <i>"Teoria da Visualização de Dados" ou "Data Visualization Theory"</i>	
<b>Informações principais sobre os dados</b>	
<b>Período:</b>	1996:2023
<b>Fontes (Jornais, Livros etc.):</b>	55
<b>Palavras-chave dos autores:</b>	48
<b>Autores:</b>	55
<b>Artigo:</b>	25
<b>Leis ou regulamentações e Documentos governamentais:</b>	2
<b>Livro e seção de Livro:</b>	5
<b>Programa de computador:</b>	1
<b>Conferências:</b>	16
<b>Teses:</b>	2

Quadro B. 12: Síntese das fontes de dados públicas brasileiras identificados geralmente em estudos canais de PD&S, a partir do referencial teórico (Seção 2)

ID	Fonte de dados Públicos*	Sigla	Aporte teórico
1	Plataforma LATTES	LATTES	(Garcia, 2011), (Perucchi <i>et al.</i> , 2017) e (Santos <i>et al.</i> , 2015)
2	Diretório do grupo de Pesquisa	DGP	(Chaves <i>et al.</i> , 2017), (Garcia, 2011) (Macedo, 2014), (Miano; Do Couto; <i>et al.</i> , 2017) e (Miano <i>et al.</i> , 2020)
3	Relatório de gestão	RGIE	(Lino, 2013) e (Bento, 2015)
4	Instituto nacional de Propriedade Industrial	INPI	(Chaves <i>et al.</i> , 2017) e (Menezes <i>et al.</i> , 2019)

\* Nota: Outras fontes (Miranda *et al.*, 2018; Niwa, 2014; Perucchi *et al.*, 2015; Ribeiro *et al.*, 2020) foram identificadas, como questionários, grupos focais, análise documental etc.

Quadro B. 13: Expansão das fontes de dados oficiais brasileiras consideradas (Quadro B. 12), incluindo rótulo e sigla.

Id	Rótulo	Sigla	Descrição
F1	Lei orçamentária anual	LOA	Estabelece as despesas e as receitas que serão realizadas pelos órgãos e autarquias da federação, dentre elas o dispêndio em pesquisa e desenvolvimento tecnológico destinados às ICTs
F2	Portal da Transparência	PTB	Apresenta informações relacionadas à administração Pública Federal do Brasil
F3	Portal do Censo Superior da Educação	CES	Apresenta um conjunto de indicadores educacionais referentes as ICTs
F4	Diretório dos Grupos de Pesquisa do CNPQ	DGP	Disponibiliza o inventário dos grupos em atividade no Brasil
F5	Plataforma Lattes	Lattes	Disponibiliza o registro de dados relacionados às atividades acadêmicas e profissionais pregressa e atual dos estudantes e pesquisadores brasileiros e estrangeiros no país
F6	Instituto Nacional de Propriedade Industrial	INPI	Disponibiliza acesso aos registros de propriedade industrial no Brasil
F7	Relatórios de gestão das ICTs da RFEPCCT	RGIEP	Disponibiliza os resultados anuais das ICTs relacionados a várias dimensões, dentre as quais ensino, pesquisa, extensão e aplicação de recursos

Quadro B. 14: Características das 28 variáveis/métricas consideradas para a coleta dos dados referentes aos canais de PD&S.

<b>Id</b>	<b>Rótulo</b>	<b>Categoria</b>	<b>Quadro B. 5</b>	<b>Tipo</b>	<b>Link TH</b>	<b>Quadro B. 13</b>	<b>Descrição</b>
<b>W<sub>1</sub></b>	CodRegiao	Geográfica	Adicional	filtro (tabela dimensão)	-	F <sub>4</sub>	Informa a localização administrativa (Regional), na qual a ICT está inserida
<b>W<sub>2</sub></b>	CodUF	Geográfica	Adicional	filtro (tabela dimensão)	-	F <sub>4</sub>	Informa a localização administrativa (unidade federativa ou Estado), na qual a ICT está inserida
<b>W<sub>3</sub></b>	CodOrgao	Geográfica	Adicional	filtro (tabela dimensão)	-	F <sub>4</sub>	Representa as siglas ICTs vinculadas à Rede Federal de Educação Profissional, Científica e Tecnológica (RFEPC) consideradas.
<b>T<sub>1</sub></b>	CodTempo	Temporal	Adicional	filtro (tabela dimensão)	-	F <sub>1</sub> ; F <sub>7</sub>	Informa o período (ANO), no qual o valor do dado está vinculado. Os dados foram coletados entre 01/01/2008 a 31/12/2019. Após processo de ETL os dados considerados para análise foram aqueles cujo período (ANO) está contido entre o intervalo 2008 e 2018.
<b>X<sub>1</sub></b>	QtdTotalPesquisadores	Institucional	ID 09	métrica/valor (tabela fato)	GOV	F <sub>1</sub> ; F <sub>3</sub> ; F <sub>7</sub>	Valor total do investimento público federal em pesquisa e desenvolvimento. Os valores foram coletados diretamente nos microdados do censo da educação superior (CES). Nesta fonte de dados, os casos identificados como atípicos (por exemplo, quando seus valores foram registrados como zero para um determinado ano) foram submetidos a uma tripla checagem. A primeira verificou os relatórios de gestão das respectivas ICTs, não sendo possível identificar tal dado foi realizada a segunda checagem. Na segunda checagem foi enviada uma consulta direta via e-mail aos canais oficiais das respectivas ICTs utilizando o serviço de <i>help desk</i> do Serviço de Informação ao Cidadão (SIC-CD) do governo federal, não havendo contestação ou atualização do respectivo dado foi mantido o valor registrado na fonte de dado primária. Na terceira a LOA foi consultada, porém como não foi



							possível identificar com precisão o total de recursos federal destinado ao fomento de pesquisa e desenvolvimento tecnológico, manteve-se o dado obtido do CES. Destaca-se que esse recurso é utilizado para o fomento nos três níveis de formação que as ICTs atuam e que no mínimo 50% dos alunos são do ensino médio técnico, que é uma imposição legal desde a criação da Rede Federal, conforme Lei nº 11.892/2008 (Brasil, 2008), e que nem todas as instituições possuíam pós-graduação. A única ICTs fora desse padrão é a UTFPR.
--	--	--	--	--	--	--	--

Continua...

Continua...

<b>Rótulo</b>	<b>Categoria</b>	<b>Quadro B. 5</b>	<b>Tipo</b>	<b>Link TH</b>	<b>Quadro B. 13</b>	<b>Descrição</b>	
<b>X2</b>	QtdTotalde pesqisadpres	Institucional	ID 19	métrica/valor (tabela fato)	GOV	F2; F4; F5	Quantidade de pesquisadores (docentes e técnicos) ativos por ICTs com pelo menos uma produção no período
<b>X3</b>	QtdTotalOrientacaoPesq	Pesquisa	ID 20	métrica/valor (tabela fato)	ICTs	F5	Quantidade de orientações de pesquisa (concluídas), desenvolvidas exclusivamente por uma ou mais ICTs. Compreende os seguintes tipos: iniciação científica; trabalho de conclusão de curso de graduação; monografia; dissertação de mestrado; tese de doutorado; orientação de pós-doutorado; outras orientações
<b>X4</b>	QtdTotalProjetPesq	Pesquisa	ID 23	métrica/valor (tabela fato)	ICTs	F5	Quantidade de projetos de pesquisa (concluídos ou em andamento), desenvolvidas exclusivamente por uma ou mais ICTs. Compreende projetos nas áreas de: desenvolvimento tecnológico, ensino, extensão, pesquisa e outros; abrangendo do ensino médio técnico à pós-graduação
<b>X5</b>	QtdTotalProjetPD	Pesquisa	ID 31	métrica/valor (tabela fato)	ICTs	F5	Quantidade de projetos de desenvolvimento tecnológico (concluídos ou em andamento), desenvolvidas exclusivamente por uma ou mais ICTs. Observação: atributo subjetivo, compreende todos os tipos de produções considerados pelos seus respectivos pesquisadores como sendo de P&D

Continua...

Continua...

	<b>Rótulo</b>	<b>Categoria</b>	<b>Quadro B. 5</b>	<b>Tipo</b>	<b>Link TH</b>	<b>Quadro B. 13</b>	<b>Descrição</b>
<b>X6</b>	QtdTotalArtigo CompletoPub	Pesquisa	ID 16	métrica/valor (tabela fato)	ICTs	F5	Quantidade de artigos completos (indexados ou não) cuja autoria e publicação por uma ou mais ICTs pode conter os seguintes atores ICTs-SPP-GOV. Neste estudo não foi possível determinar a origem dos atores da TH, pois não há nesta base um marcador objetivo para tal finalidade. Entretanto, em cada artigo contabilizado há pelo menos um autor vinculado a uma ou mais ICTs. Para efeito de contagem foi considerado apenas um registro do título do artigo por ICTs. Obs. a quantidade total geral contém contagem duplicada para artigos publicados cuja autoria e publicação ocorreu por mais de uma ICTs no mesmo período
<b>X7</b>	QtdTotalOrientacaoPesqFinancSPP	Pesquisa	Adicional	métrica/valor (tabela fato)	ICTs - SPP	F5	Quantidade de orientações de pesquisa financiadas pelos setores produtivos públicos e privados (SPP)
<b>X8</b>	QtdTotalProjPesqCoopSPP	Pesquisa	Adicional	métrica/valor (tabela fato)	ICTs - SPP	F5	Quantidade de projetos de pesquisa em cooperação com setores produtivos públicos e privados (SPP)
<b>X9</b>	QtdTotalProjPesqFinancSPP	Pesquisa	Adicional	métrica/valor (tabela fato)	ICTs - SPP	F5	Quantidade de projetos de pesquisa financiados pelos setores produtivos públicos e privados (SPP)

Continua...

Continua...

	<b>Rótulo</b>	<b>Categoria</b>	<b>Quadro B. 5</b>	<b>Tipo</b>	<b>Link TH</b>	<b>Quadro B. 13</b>	<b>Descrição</b>
<b>X10</b>	QtdTotalTreinamentos	prestação de serviço técnicos especializados	ID 43	métrica/valor (tabela fato)	ICTs	F5	Quantidade de treinamentos demandados
<b>X11</b>	QtdTotalConsultorias	prestação de serviço técnicos especializados	ID 39	métrica/valor (tabela fato)	ICTs	F5	Quantidade de consultorias demandadas
<b>X12</b>	QtdTotalExt	prestação de serviço técnicos especializados	Adicional	métrica/valor (tabela fato)	ICTs	F5	Quantidade de projetos de extensão (acadêmicos ou tecnológicos). Observação: atributo subjetivo, pois não a na base um marcador que permita objetivamente selecionar uma das duas opções
<b>X13</b>	QtdTotalCessaoInfraestrutura	prestação de serviço técnicos especializados	Adicional	métrica/valor (tabela fato)	ICTs	F5; F4	Quantidade de serviços de uso de laboratórios e equipamentos; etc.
<b>X14</b>	QtdTotalTreinamentosSPP	prestação de serviço técnicos especializados	Adicional	métrica/valor (tabela fato)	ICTs - SPP	F5	Quantidade de treinamentos demandados pelos setores produtivos públicos e privados (SPP)

Continua...

Continua...

	<b>Rótulo</b>	<b>Categoria</b>	<b>Quadro B. 5</b>	<b>Tipo</b>	<b>Link TH</b>	<b>Quadro B. 13</b>	<b>Descrição</b>
X15	QtdTotalConsultoriasSPP	prestação de serviço técnicos especializados	ID 40	métrica/valor (tabela fato)	ICTs - SPP	F5	Quantidade de consultorias demandadas pelos setores produtivos públicos e privados (SPP)
X16	QtdTotalExtSP	prestação de serviço técnicos especializados	Adicional	métrica/valor (tabela fato)	ICTs - SPP	F5	Quantidade de projetos de extensão (acadêmicos ou tecnológicos) demandados pelos setores produtivos públicos e privados (SPP). Observação: atributo subjetivo, pois não a na base um marcador que permita objetivamente selecionar uma das duas opções
X17	QtdTotalCessaoInfraestruturaSPP	prestação de serviço técnicos especializados	Adicional	métrica/valor (tabela fato)	ICTs - SPP	F5; F4	Quantidade de serviços de uso de laboratórios e equipamentos; etc., demandados pelos setores produtivos públicos e privados (SPP)
X18	QtdTotalMarcas	Desenvolvimento tecnológico	Adicional	métrica/valor (tabela fato)	ICTs	F6	Quantidade de marcas (concedidas), compreendendo marcas comerciais e institucionais cuja titularidade é exclusiva de uma ou mais ICTs
X19	QtdTotalDesIn	Desenvolvimento tecnológico	Adicional	métrica/valor (tabela fato)	ICTs	F6	Quantidade de desenhos industriais (concedidos) cuja titularidade é exclusiva de uma ou mais ICTs

Continua...

ontinua...

Rótulo		Categoria	Quadro B. 5	Tipo	Link TH	Quadro B. 13	Descrição
X20	QtdTotalPatentes	Desenvolvimento tecnológico	ID 10	métrica/valor (tabela fato)	ICTs	F <sub>6</sub>	Quantidade de patentes (publicadas) cuja titularidade é exclusiva de uma ou mais ICTs
X21	QtdTotalProgComputadores	Desenvolvimento tecnológico	Adicional	métrica/valor (tabela fato)	ICTs	F <sub>6</sub>	Quantidade de programas de computador (registrados) cuja titularidade é exclusiva de uma ou mais ICTs
X22	QtdTotalDesIndsSPP	Desenvolvimento tecnológico	Adicional	métrica/valor (tabela fato)	ICTs - SPP	F <sub>6</sub>	Quantidade de desenhos industriais (concedidos) cuja titularidade é de uma ou mais ICTs em parceria com pelo menos um autor representante dos setores públicos e privados (SPP)
X23	QtdTotalPatentesSPP	Desenvolvimento tecnológico	ID 11	métrica/valor (tabela fato)	ICTs - SPP	F <sub>6</sub>	Quantidade de patentes (publicadas) cuja titularidade é de uma ou mais ICTs em parceria com pelo menos um autor representante dos setores públicos e privados (SPP)
X24	QtdTotalProgComputadoresSPP	Desenvolvimento tecnológico	Adicional	métrica/valor (tabela fato)	ICTs - SPP	F <sub>6</sub>	Quantidade de programas de computador (registrados) cuja titularidade é de uma ou mais ICTs em parceria com pelo menos um autor representante dos setores públicos e privados (SPP)

Quadro B. 15: Resumo das Variáveis Utilizadas e Suas Categorizações no Estudo

<b>Id</b>	<b>Nome</b>	<b>Categoria</b>	<b>Setor Triple Helix</b>
X <sub>1</sub>	Investimento	Inst.	GOV
X <sub>2</sub>	Pesquisadores	Inst.	GOV
X <sub>3</sub>	Orientação e Pesquisa	Pesq.	ICTs
X <sub>4</sub>	Projetos de Pesquisa	Pesq.	ICTs
X <sub>5</sub>	Projeto PD	Pesq.	ICTs
X <sub>6</sub>	Publicação	Pesq.	ICTs
X <sub>7</sub>	Orientação de Pesquisa SPP	Pesq.	ICTs -SPP
X <sub>8</sub>	Projetos de Pesquisa Cooperação SPP	Pesq.	ICTs -SPP
X <sub>9</sub>	Projeto de Pesquisa Financiada SPP	Pesq.	ICTs -SPP
X <sub>10</sub>	Treinamento	Serv.	ICTs
X <sub>11</sub>	Consultoria	Serv.	ICTs
X <sub>12</sub>	Extensão Acadêmica	Serv.	ICTs
X <sub>13</sub>	Outras atividades	Serv.	ICTs
X <sub>14</sub>	Treinamento SPP	Serv.	ICTs -SPP
X <sub>15</sub>	Consultoria SPP	Serv.	ICTs -SPP
X <sub>16</sub>	Projetos de Extensão SPP	Serv.	ICTs -SPP
X <sub>17</sub>	Outros Serviços Tec SPP	Serv.	ICTs -SPP
X <sub>18</sub>	Marcas	Des. Tec.	ICTs
X <sub>19</sub>	Desenho Industrial	Des. Tec.	ICTs
X <sub>20</sub>	Patentes	Des. Tec.	ICTs
X <sub>21</sub>	Programa de Computador	Des. Tec.	ICTs
X <sub>22</sub>	Desenho Industrial SPP	Des. Tec.	ICTs -SPP
X <sub>23</sub>	Patentes SPP	Des. Tec.	ICTs -SPP
X <sub>24</sub>	Programa de Computadores SPP	Des. Tec.	ICTs -SPP

Nota: Na análise de agrupamentos realizada, foram consideradas duas abordagens distintas quanto à inclusão de variáveis. Abordagem 1 considerou todas as 24 variáveis do estudo, independentemente de suas médias ( $\mu$ ) e desvios padrão ( $\sigma$ ), mesmo que apresentassem valores iguais a zero. Abordagem 2: Incluiu apenas variáveis com médias ( $\mu$ ) e desvios padrão ( $\sigma$ ) diferentes de zero e aquelas identificadas como “contributivas” após a análise ANOVA.

Tabela B. 3: Somatório dos dados referentes aos canais de PD&S das 42 unidades da Rede Federal Brasileira de Educação Profissional, Científica e Tecnológica entre 2008 e 2018. A descrição dos dados está disponível no Apêndice B, *Quadro B. 14*: Região ( $W_1$ ), Região ( $UF_2$ ), Órgão ( $W_3$ ), Investimento público em Pesquisa e Desenvolvimento ( $X_1$ ), Pesquisadores ( $X_2$ ), Orientações de Pesquisa ( $X_3$ ), Projetos de Pesquisa ( $X_4$ ), Linhas de Pesquisa ( $X_5$ ), Publicações científicas ( $X_6$ ), Pesquisas financiadas pelos setores produtivos ( $X_7$ ), Linhas de pesquisa com setores produtivos ( $X_8$ ), Projetos de Pesquisas financiadas pelos setores produtivos ( $X_9$ ), Treinamento ( $X_{10}$ ), Consultoria ( $X_{11}$ ), Projetos de extensão acadêmicos ou tecnológicos ( $X_{12}$ ), Serviços de uso de laboratórios e equipamentos; etc. ( $X_{13}$ ), Treinamentos demandados pelos setores produtivos ( $X_{14}$ ), Consultorias demandadas pelos setores produtivos ( $X_{15}$ ), Projetos de extensão demandados pelos setores produtivos ( $X_{16}$ ), Serviços de uso de laboratórios e equipamentos; etc. demandados pelos setores produtivos ( $X_{17}$ ), Marcas (concedidas) ( $X_{18}$ ), Desenhos Industriais (concedidos) ( $X_{19}$ ), Patentes (publicadas) ( $X_{20}$ ), Programas de computador (concedidos) ( $X_{21}$ ), Desenhos Industriais (concedidos) demandados pelos setores produtivos ( $X_{22}$ ), Patentes (publicadas) demandadas pelos setores produtivos ( $X_{23}$ ) e Programas de computador (concedidos) demandados pelos setores produtivos ( $X_{24}$ ).

$W_3$	$W_1$	$W_2$		$X_1$	$X_2$	$X_3$	$X_4$	$X_5$	$X_6$
<b>CEFET/MG</b>	SW	MG	R\$	5.432.870.222,00	929	10774	2998	480	2865
<b>CEFET/RJ</b>	SW	RJ	R\$	633.394.355,00	728	4955	1796	313	2409
<b>CPII</b>	SW	RJ	R\$	0,00	740	2068	1025	278	1278
<b>IFAC</b>	N	AC	R\$	1.259.585,00	322	887	957	162	357
<b>IFAL</b>	NW	AL	R\$	17.354.220,00	844	4096	2964	361	1323
<b>IFAM</b>	N	AM	R\$	378.045.352,00	818	4344	2574	234	1147
<b>IFAP</b>	N	AP	R\$	136.765.644,00	287	783	805	71	207
<b>IFB</b>	CW	DF	R\$	693.106.825,00	611	1974	1403	321	858
<b>IFBA</b>	NW	BA	R\$	588.007.349,00	1257	6850	3627	710	2024
<b>IFBAIANO</b>	NW	BA	R\$	372.999.607,00	802	3396	2130	299	1516
<b>IFC</b>	S	SC	R\$	280.984.037,00	1014	7636	4036	583	2193
<b>IFCE</b>	NW	CE	R\$	1.551.945.731,00	1575	13108	5336	685	4181
<b>IFES</b>	SW	ES	R\$	5.619.408.098,00	1462	11106	4883	820	3927
<b>IFF</b>	SW	RJ	R\$	536.082.283,00	852	3340	2236	304	1821
<b>IFFAR</b>	S	RS	R\$	357.297.926,00	893	6254	4313	389	1897

Continua...



Continua...

<b>W<sub>3</sub></b>		<b>W<sub>1</sub></b>	<b>W<sub>2</sub></b>	<b>X<sub>1</sub></b>	<b>X<sub>2</sub></b>	<b>X<sub>3</sub></b>	<b>X<sub>4</sub></b>	<b>X<sub>5</sub></b>	<b>X<sub>6</sub></b>	<b>X<sub>1</sub></b>	<b>X<sub>2</sub></b>	<b>X<sub>3</sub></b>	<b>X<sub>4</sub></b>
<b>IFGOIANO</b>	CW	GO	R\$	1.257.672.877,00	787		7974		4004		496		4148
<b>IFMA</b>	NW	MA	R\$	1.714.748.562,00	1537		8268		6797		623		2230
<b>IFMG</b>	SW	MG	R\$	1.595.868.171,00	938		6116		2709		360		1522
<b>IFMS</b>	CW	MS	R\$	249.639.037,00	558		2487		2066		244		803
<b>IFMT</b>	CW	MT	R\$	897.359.926,00	1000		6658		3963		344		2191
<b>IFNMG</b>	SW	MG	R\$	244.572.857,00	704		3548		2712		300		1145
<b>IFPA</b>	N	PA	R\$	3.642.612.226,00	1183		4875		3681		466		1823
<b>IFPB</b>	NW	PB	R\$	8.160.175.303,00	1181		7771		4187		474		2359
<b>IFPE</b>	NW	PE	R\$	117.056.226,00	1016		5589		2333		429		1684
<b>IFPI</b>	NW	PI	R\$	1.446.509.716,00	1138		6025		3394		345		1734
<b>IFPR</b>	S	PR	R\$	657.736.605,00	1250		6339		4930		646		2618
<b>IFRJ</b>	SW	RJ	R\$	501.453.714,00	1016		7781		3773		679		2687
<b>IFRN</b>	NW	RN	R\$	1.448.556.552,00	1527		11687		6261		674		3088
<b>IFRO</b>	N	RO	R\$	374.584.206,00	670		2926		3308		310		1190
<b>IFRR</b>	N	RR	R\$	485.109.424,00	269		1478		958		50		291
<b>IFRS</b>	S	RS	R\$	1.200.531.707,00	1333		9049		7367		808		2676
<b>IFSC</b>	S	SC	R\$	836.423.912,00	1552		6380		6837		905		3096
<b>IFSE</b>	NW	SE	R\$	1.295.085.213,00	519		2455		1828		267		919
<b>IFSEMG</b>	SW	MG	R\$	518.790.766,00	647		5895		3231		265		1448
<b>IFSMG</b>	SW	MG	R\$	1.230.227.058,00	607		4830		2603		271		1506
<b>IFSP</b>	SW	SP	R\$	139.513.959,00	2298		11048		4533		1318		4180
<b>IFSTPE</b>	NW	PE	R\$	1.385.054.456,00	457		2277		1934		209		967
<b>IFSUL</b>	S	RS	R\$	604.992.797,00	945		7004		3499		342		1798
<b>IFTM</b>	SW	MG	R\$	265.478.829,00	644		4961		4122		372		1545
<b>IFTO</b>	N	TO	R\$	41.942.126,00	642		3078		2239		264		1161
<b>UTFPR</b>	S	PR	R\$	6.709.975.942,00	2656		50110		12108		1862		16445

Continua...

Continua...

<b>W<sub>3</sub></b>	<b>W<sub>1</sub></b>	<b>W<sub>2</sub></b>	<b>X<sub>7</sub></b>	<b>X<sub>8</sub></b>	<b>X<sub>9</sub></b>	<b>X<sub>10</sub></b>	<b>X<sub>11</sub></b>	<b>X<sub>12</sub></b>	<b>X<sub>13</sub></b>	<b>X<sub>14</sub></b>
<b>CEFET/MG</b>	SW	MG	427	0	57	24	61	119	156	0
<b>CEFET/RJ</b>	SW	RJ	106	5	10	13	44	121	54	0
<b>CPII</b>	SW	RJ	189	1	5	8	40	17	54	0
<b>IFAC</b>	N	AC	17	0	3	11	25	30	32	0
<b>IFAL</b>	NW	AL	22	0	21	17	45	65	128	1
<b>IFAM</b>	N	AM	224	0	9	19	51	52	120	0
<b>IFAP</b>	N	AP	10	0	6	18	54	23	36	0
<b>IFB</b>	CW	DF	54	0	7	17	57	58	58	0
<b>IFBA</b>	NW	BA	137	0	29	24	96	162	173	0
<b>IFBAIANO</b>	NW	BA	128	0	20	17	86	83	191	1
<b>IFC</b>	S	SC	12	1	17	16	189	205	263	0
<b>IFCE</b>	NW	CE	501	1	206	33	134	218	266	0
<b>IFES</b>	SW	ES	209	4	122	48	137	200	228	0
<b>IFF</b>	SW	RJ	9	1	26	14	75	65	56	0
<b>IFFAR</b>	S	RS	6	0	4	28	217	138	261	0
<b>IFGO</b>	CW	GO	35	2	35	20	99	136	162	0
<b>IFGOIANO</b>	CW	GO	20	0	20	7	76	114	167	0
<b>IFMA</b>	NW	MA	344	0	19	25	99	129	203	0
<b>IFMG</b>	SW	MG	29	0	15	32	143	176	151	0
<b>IFMS</b>	CW	MS	12	1	5	7	57	53	68	0
<b>IFMT</b>	CW	MT	41	0	16	32	111	73	123	0
<b>IFNMG</b>	SW	MG	62	0	14	15	67	95	80	0
<b>IFPA</b>	N	PA	148	0	15	23	160	168	203	0
<b>IFPB</b>	NW	PB	55	0	14	11	88	185	137	0
<b>IFPE</b>	NW	PE	76	0	18	55	154	154	182	0

Continua...

Continua...

<b>W<sub>3</sub></b>		<b>W<sub>1</sub></b>	<b>W<sub>2</sub></b>	<b>X<sub>7</sub></b>	<b>X<sub>8</sub></b>	<b>X<sub>9</sub></b>	<b>X<sub>10</sub></b>	<b>X<sub>11</sub></b>	<b>X<sub>12</sub></b>	<b>X<sub>13</sub></b>	<b>X<sub>14</sub></b>	<b>X<sub>7</sub></b>	<b>X<sub>8</sub></b>
<b>IFPI</b>	NW	PI	32	0	8	14	82	118	69	0			
<b>IFPR</b>	S	PR	14	0	9	13	106	182	227	0			
<b>IFRJ</b>	SW	RJ	281	1	63	17	55	187	176	0			
<b>IFRN</b>	NW	RN	342	3	61	41	103	193	187	0			
<b>IFRO</b>	N	RO	15	2	5	21	90	67	79	0			
<b>IFRR</b>	N	RR	20	1	1	2	15	5	29	0			
<b>IFRS</b>	S	RS	101	3	18	36	149	493	335	0			
<b>IFSC</b>	S	SC	65	6	49	35	150	451	271	0			
<b>IFSE</b>	NW	SE	60	4	15	13	73	30	60	0			
<b>IFSEMG</b>	SW	MG	116	1	14	20	80	99	139	0			
<b>IFSMG</b>	SW	MG	28	0	15	20	53	90	123	0			
<b>IFSP</b>	SW	SP	162	1	20	31	192	659	331	0			
<b>IFSTPE</b>	NW	PE	4	0	5	13	37	50	71	0			
<b>IFSUL</b>	S	RS	14	2	20	28	130	112	115	0			
<b>IFTM</b>	SW	MG	8	0	22	25	78	114	130	1			
<b>IFTO</b>	N	TO	45	0	5	13	111	66	84	0			
<b>UTFPR</b>	S	PR	363	3	266	99	276	774	704	0			

Continua...

Continua...

<b>W<sub>3</sub></b>	<b>W<sub>1</sub></b>	<b>W<sub>2</sub></b>	<b>X<sub>15</sub></b>	<b>X<sub>16</sub></b>	<b>X<sub>17</sub></b>	<b>X<sub>18</sub></b>	<b>X<sub>19</sub></b>	<b>X<sub>20</sub></b>	<b>X<sub>21</sub></b>	<b>X<sub>22</sub></b>	<b>X<sub>23</sub></b>	<b>X<sub>24</sub></b>
<b>CEFET/MG</b>	SW	MG	0	0	0	31	0	35	38	0	1	1
<b>CEFET/RJ</b>	SW	RJ	0	2	0	0	0	2	0	0	0	0
<b>CPII</b>	SW	RJ	0	0	0	0	0	0	0	0	0	0
<b>IFAC</b>	N	AC	0	0	0	0	0	1	0	0	0	0
<b>IFAL</b>	NW	AL	0	0	0	0	0	0	5	0	0	0
<b>IFAM</b>	N	AM	0	0	0	0	0	6	0	0	0	0
<b>IFAP</b>	N	AP	0	0	0	0	0	0	0	0	0	0
<b>IFB</b>	CW	DF	0	0	0	0	0	0	3	0	0	0
<b>IFBA</b>	NW	BA	2	0	0	4	1	26	13	0	0	0
<b>IFBAIANO</b>	NW	BA	0	0	0	0	0	10	1	0	0	0
<b>IFC</b>	S	SC	0	0	0	0	0	2	2	0	0	0
<b>IFCE</b>	NW	CE	1	5	0	0	0	28	5	0	0	0
<b>IFES</b>	SW	ES	0	0	0	9	2	50	14	0	1	0
<b>IFF</b>	SW	RJ	0	1	0	0	0	0	4	0	0	0
<b>IFFAR</b>	S	RS	0	0	0	7	2	2	4	0	0	0
<b>IFGO</b>	CW	GO	0	0	0	4	0	6	12	0	0	0
<b>IFGOIANO</b>	CW	GO	1	0	0	0	0	1	3	0	0	0
<b>IFMA</b>	NW	MA	0	0	0	0	0	7	0	0	0	0
<b>IFMG</b>	SW	MG	0	0	0	3	0	4	23	0	0	0
<b>IFMS</b>	CW	MS	0	0	0	0	0	1	2	0	0	0
<b>IFMT</b>	CW	MT	0	1	0	0	0	3	4	0	0	0
<b>IFNMG</b>	SW	MG	0	0	0	1	0	0	0	0	0	0
<b>IFPA</b>	N	PA	2	1	1	1	0	3	0	0	0	0
<b>IFPB</b>	NW	PB	3	1	0	2	0	19	18	0	0	0
<b>IFPE</b>	NW	PE	4	1	0	0	2	4	4	0	0	0
<b>IFPI</b>	NW	PI	1	0	0	0	0	9	4	0	0	0

Continua...

Continua...

<b>W<sub>3</sub></b>	<b>W<sub>1</sub></b>	<b>W<sub>2</sub></b>	<b>X<sub>15</sub></b>	<b>X<sub>16</sub></b>	<b>X<sub>17</sub></b>	<b>X<sub>18</sub></b>	<b>X<sub>19</sub></b>	<b>X<sub>20</sub></b>	<b>X<sub>21</sub></b>	<b>X<sub>22</sub></b>	<b>X<sub>23</sub></b>	<b>X<sub>24</sub></b>
<b>IFPR</b>	S	PR	0	1	0	0	0	0	0	0	0	0
<b>IFRJ</b>	SW	RJ	0	0	0	0	0	3	1	0	0	0
<b>IFRN</b>	NW	RN	0	0	0	0	0	10	56	0	0	0
<b>IFRO</b>	N	RO	0	0	0	0	0	9	2	0	0	0
<b>IFRR</b>	N	RR	0	0	0	0	0	0	0	0	0	0
<b>IFRS</b>	S	RS	0	0	0	0	0	1	7	0	1	0
<b>IFSC</b>	S	SC	0	0	0	0	7	25	4	1	1	2
<b>IFSE</b>	NW	SE	0	0	1	5	0	12	12	0	0	0
<b>IFSEMG</b>	SW	MG	0	0	0	2	2	6	8	0	0	0
<b>IFSMG</b>	SW	MG	0	1	0	3	0	4	9	0	1	0
<b>IFSP</b>	SW	SP	0	0	0	2	0	11	3	0	0	0
<b>IFSTPE</b>	NW	PE	0	0	0	0	0	8	3	0	0	0
<b>IFSUL</b>	S	RS	1	0	0	0	0	10	9	0	0	0
<b>IFTM</b>	SW	MG	0	0	0	0	0	1	1	0	0	0
<b>IFTO</b>	N	TO	0	0	0	0	0	0	0	0	0	0
<b>UTFPR</b>	S	PR	0	4	0	9	0	108	34	0	4	2

Tabela B. 4: Estatísticas descritivas dos dados referentes aos canais de PD&S das 42 unidades da Rede Federal Brasileira de Educação Profissional, Científica e Tecnológica entre 2008 e 2018

<b>Estatística</b>	<b>X1</b>	<b>X2</b>	<b>X3</b>	<b>X4</b>	<b>X5</b>	<b>X6</b>	<b>X7</b>	<b>X8</b>	<b>X9</b>	<b>X10</b>	<b>X11</b>	<b>...</b>
<b>Sum</b>	5,66E+14	41360.00	284582.00	149842.00	19928.00	95440.00	4543.00	43.00	1309.00	975.00	4145.00	...
<b>Std_Dev</b>	1,84E+13	484.28	7491.57	2096.54	327.96	2456.40	127.78	1.54	51.68	16.31	54.81	...
<b>Std_Error</b>	2,84E+12	74.73	1155.97	323.50	50.61	379.03	19.72	0.24	7.97	2.52	8.46	...
<b>Variance</b>	3,39E+20	234523.55	56123602.19	4395464.03	107559.91	6033879.56	16328.58	2.37	2670.48	265.93	3004.66	...
<b>Coeff_Variation</b>	136,62	49.18	110.56	58.76	69.12	108.10	118.14	150.22	165.81	70.25	55.54	...
<b>IQR</b>	1,06E+13	529.75	4383.25	1934.00	332.75	1353.75	127.50	1.00	13.50	14.75	76.00	...
<b>CI_Low</b>	7,74E+12	833.85	4441.22	2914.34	372.28	1506.91	68.35	0.54	15.06	18.13	81.61	...
<b>CI_High</b>	1,92E+13	1135.67	9110.30	4220.99	576.68	3037.85	147.99	1.50	47.27	28.30	115.77	...
<b>Mode</b>	0	1016.00	783.00	805.00	50.00	207.00	12.00	0.00	5.00	13.00	57.00	...
<b>Mean</b>	1,35E+13	984.76	6775.76	3567.67	474.48	2272.38	108.17	1.02	31.17	23.21	98.69	...
<b>Median</b>	6,46E+12	911.00	5960.00	3345.00	360.50	1809.50	54.50	0.00	15.50	19.50	87.00	...

Continua...

<b>Estatística</b>	<b>X12</b>	<b>X13</b>	<b>X14</b>	<b>X15</b>	<b>X16</b>	<b>X17</b>	<b>X18</b>	<b>X19</b>	<b>X20</b>	<b>X21</b>	<b>X22</b>	<b>X23</b>	<b>X24</b>
<b>Sum</b>	6529.00	6682.00	3.00	15.00	18.00	2.00	83.00	16.00	427.00	308.00	1.00	9.00	5.00
<b>Std_Dev</b>	159.96	118.01	0.26	0.88	1.04	0.22	5.17	1.21	18.83	11.54	0.15	0.68	0.45
<b>Std_Error</b>	24.68	18.21	0.04	0.14	0.16	0.03	0.80	0.19	2.91	1.78	0.02	0.11	0.07
<b>Variance</b>	25586.25	13926.04	0.07	0.77	1.08	0.05	26.76	1.46	354.48	133.06	0.02	0.47	0.20
<b>Coeff_Variation</b>	102.90	74.17	364.93	245.98	242.50	452.63	261.74	317.30	185.19	157.30	648.07	318.28	380.32
<b>IQR</b>	115.25	127.00	0.00	0.00	0.00	0.00	2.00	0.00	9.00	8.50	0.00	0.00	0.00
<b>CI_Low</b>	105.61	122.32	-0.01	0.08	0.10	-0.02	0.36	0.00	4.30	3.74	-0.02	0.00	-0.02
<b>CI_High</b>	205.30	195.87	0.15	0.63	0.75	0.11	3.59	0.76	16.03	10.93	0.07	0.43	0.26
<b>Mode</b>	30.00	54.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
<b>Mean</b>	155.45	159.10	0.07	0.36	0.43	0.05	1.98	0.38	10.17	7.33	0.02	0.21	0.12
<b>Median</b>	116.00	138.00	0.00	0.00	0.00	0.00	0.00	0.00	4.00	3.50	0.00	0.00	0.00

Tabela B. 5: Análise das correções dos dados referentes as 24 variáveis relacionada aos canais de PD&S das 41 unidades da Rede Federal  
 Análise das correções dos dados referentes as 24 variáveis relacionada aos canais de PD&S das 41 unidades da Rede Federal Brasileira de  
 Educação Profissional, Científica e Tecnológica (outlier: UTFPR), entre 2008 e 2018

	X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>	X <sub>4</sub>	X <sub>5</sub>	X <sub>6</sub>	X <sub>7</sub>	X <sub>8</sub>	X <sub>9</sub>	X <sub>10</sub>
X <sub>1</sub>	1									
X <sub>2</sub>	0.25 (0.11)	1								
X <sub>3</sub>	0.42 ** (0.01)	0.82 *** (0)	1							
X <sub>4</sub>	0.24 (0.13)	0.78 *** (0)	0.77 *** (0)	1						
X <sub>5</sub>	0.19 (0.24)	0.93 *** (0)	0.78 *** (0)	0.76 *** (0)	1					
X <sub>6</sub>	0.34 * (0.03)	0.81 *** (0)	0.9 *** (0)	0.72 *** (0)	0.83 *** (0)	1				
X <sub>7</sub>	0.3 . (0.06)	0.49 ** (0)	0.64 *** (0)	0.37 * (0.02)	0.4 ** (0.01)	0.51 *** (0)	1			
X <sub>8</sub>	0.03 (0.86)	0.25 (0.12)	0.2 (0.2)	0.31 * (0.05)	0.34 * (0.03)	0.31 * (0.05)	0.09 (0.57)	1		
X <sub>9</sub>	0.3 . (0.06)	0.47 ** (0)	0.67 *** (0)	0.41 ** (0.01)	0.45 ** (0)	0.64 *** (0)	0.69 *** (0)	0.27 . (0.09)	1	
X <sub>10</sub>	0.14 (0.39)	0.58 *** (0)	0.59 *** (0)	0.55 *** (0)	0.53 *** (0)	0.46 ** (0)	0.34 * (0.03)	0.27 . (0.09)	0.46 ** (0)	1
X <sub>11</sub>	0.06 (0.71)	0.65 *** (0)	0.55 *** (0)	0.59 *** (0)	0.62 *** (0)	0.51 *** (0)	0.05 (0.74)	0.14 (0.37)	0.21 (0.18)	0.63 *** (0)
X <sub>12</sub>	0.09 (0.57)	0.81 *** (0)	0.63 *** (0)	0.69 *** (0)	0.89 *** (0)	0.68 *** (0)	0.21 (0.19)	0.36 * (0.02)	0.25 (0.11)	0.5 *** (0)
X <sub>13</sub>	0.15 (0.35)	0.8 *** (0)	0.76 *** (0)	0.81 *** (0)	0.84 *** (0)	0.74 *** (0)	0.35 * (0.02)	0.18 (0.27)	0.42 ** (0.01)	0.59 *** (0)
X <sub>14</sub>	-0.07 (0.65)	-0.13 (0.44)	-0.15 (0.37)	-0.05 (0.75)	-0.11 (0.49)	-0.13 (0.42)	-0.11 (0.48)	-0.18 (0.25)	-0.03 (0.83)	-0.04 (0.79)

Continua...

Continua...

	X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>	X <sub>4</sub>	X <sub>5</sub>	X <sub>6</sub>	X <sub>7</sub>	X <sub>8</sub>	X <sub>9</sub>	X <sub>10</sub>
X <sub>15</sub>	0.34 * (0.03)	0.19 (0.23)	0.16 (0.32)	0.04 (0.8)	0.09 (0.6)	0.12 (0.44)	0.02 (0.91)	-0.21 (0.18)	0.06 (0.72)	0.27 . (0.08)
X <sub>16</sub>	0.12 (0.44)	0.22 (0.17)	0.32 * (0.04)	0.13 (0.44)	0.09 (0.56)	0.35 * (0.02)	0.38 * (0.01)	0.04 (0.79)	0.63 *** (0)	0.15 (0.35)
X <sub>17</sub>	0.17 (0.28)	-0.05 (0.75)	-0.15 (0.34)	-0.09 (0.6)	-0.07 (0.67)	-0.12 (0.44)	0 (0.98)	0.15 (0.34)	-0.07 (0.68)	-0.07 (0.67)
X <sub>18</sub>	0.51 *** (0)	0.06 (0.69)	0.33 * (0.03)	0.01 (0.97)	0.11 (0.48)	0.21 (0.18)	0.38 * (0.01)	-0.01 (0.93)	0.21 (0.19)	0.15 (0.34)
X <sub>19</sub>	0 (0.99)	0.25 (0.11)	0.12 (0.46)	0.34 * (0.03)	0.31 * (0.04)	0.22 (0.17)	-0.04 (0.82)	0.49 ** (0)	0.16 (0.32)	0.42 ** (0.01)
X <sub>20</sub>	0.61 *** (0)	0.46 ** (0)	0.58 *** (0)	0.35 * (0.03)	0.46 ** (0)	0.53 *** (0)	0.52 *** (0)	0.35 * (0.03)	0.67 *** (0)	0.45 ** (0)
X <sub>21</sub>	0.45 ** (0)	0.27 . (0.09)	0.51 *** (0)	0.27 . (0.09)	0.22 (0.18)	0.3 . (0.06)	0.41 ** (0.01)	0.18 (0.25)	0.28 . (0.08)	0.37 * (0.02)
X <sub>22</sub>	-0.04 (0.82)	0.24 (0.14)	0.03 (0.83)	0.34 * (0.03)	0.3 . (0.06)	0.18 (0.25)	-0.05 (0.76)	0.53 *** (0)	0.1 (0.52)	0.19 (0.22)
X <sub>23</sub>	0.38 * (0.02)	0.21 (0.18)	0.33 * (0.03)	0.37 * (0.02)	0.33 * (0.03)	0.33 * (0.04)	0.2 (0.22)	0.4 ** (0.01)	0.28 . (0.08)	0.38 * (0.01)
X <sub>24</sub>	0.15 (0.35)	0.21 (0.18)	0.15 (0.35)	0.29 . (0.06)	0.28 . (0.07)	0.23 (0.14)	0.15 (0.36)	0.43 ** (0)	0.16 (0.33)	0.19 (0.23)

Continua...



Continua...

	X11	X12	X13	X14	X15	X16	X17	X18	X19	X20
X11	1									
X12	0.65 *** (0)	1								
X13	0.79 *** (0)	0.82 *** (0)	1							
X14	-0.15 (0.36)	-0.12 (0.46)	0.01 (0.93)	1						
X15	0.24 (0.14)	0.06 (0.69)	0.11 (0.49)	-0.12 (0.47)	1					
X16	0.11 (0.5)	0.06 (0.7)	0.16 (0.33)	-0.11 (0.49)	0.28 . (0.07)	1				
X17	0.11 (0.51)	-0.07 (0.65)	-0.04 (0.8)	-0.06 (0.69)	0.16 (0.31)	0.04 (0.8)	1			
X18	0.03 (0.83)	0.01 (0.93)	0.11 (0.48)	-0.1 (0.53)	-0.06 (0.7)	-0.11 (0.51)	0.05 (0.74)	1		
X19	0.35 * (0.03)	0.36 * (0.02)	0.34 * (0.03)	-0.09 (0.57)	0.1 (0.55)	-0.08 (0.62)	-0.07 (0.65)	0.04 (0.78)	1	
X20	0.21 (0.19)	0.28 . (0.08)	0.34 * (0.03)	-0.11 (0.5)	0.16 (0.31)	0.18 (0.27)	-0.01 (0.97)	0.57 *** (0)	0.38 * (0.02)	1
X21	0.08 (0.61)	0.11 (0.48)	0.14 (0.37)	-0.11 (0.48)	0.04 (0.79)	-0.08 (0.63)	-0.01 (0.93)	0.52 *** (0)	-0.01 (0.94)	0.43 ** (0)
X22	0.19 (0.24)	0.39 * (0.01)	0.25 (0.12)	-0.04 (0.78)	-0.07 (0.68)	-0.06 (0.7)	-0.04 (0.82)	-0.06 (0.73)	0.87 *** (0)	0.25 (0.11)
X23	0.12 (0.44)	0.38 * (0.01)	0.35 * (0.02)	-0.1 (0.51)	-0.16 (0.33)	-0.06 (0.71)	-0.08 (0.6)	0.5 *** (0)	0.44 ** (0)	0.53 *** (0)
X24	0.12 (0.46)	0.34 * (0.03)	0.23 (0.15)	-0.06 (0.71)	-0.09 (0.58)	-0.08 (0.6)	-0.05 (0.76)	0.36 * (0.02)	0.76 *** (0)	0.41 ** (0.01)

Continua...

	X21	X22	X23	X24
X21	1			
X22	-0.04 (0.81)	1		
X23	0.27 . (0.09)	0.42 ** (0.01)	1	
X24	0.17 (0.28)	0.89 *** (0)	0.57 *** (0)	1

Tabela B. 6: Análise dos Principais Componentes das Variáveis X<sub>1</sub> a X<sub>24</sub>

<b>Variáveis</b>	<b>RPC<sub>1</sub></b>	<b>RPC<sub>2</sub></b>	<b>RPC<sub>3</sub></b>	<b>RPC<sub>4</sub></b>	<b>RPC<sub>5</sub></b>	<b>RPC<sub>6</sub></b>	<b>h<sub>2</sub></b>	<b>u<sub>2</sub></b>	<b>com</b>
<b>X<sub>1</sub> - Investimento público em Pesquisa e Desenvolvimento</b>	0,09	0	0,73	0,15	0,34	0,13	0,7	0,3	1,6
<b>X<sub>2</sub> - Pesquisadores</b>	0,9	0,06	0,12	0,23	0,04	0,05	0,88	0,117	1,2
<b>X<sub>3</sub> - Orientações de Pesquisa</b>	0,78	-0,06	0,41	0,4	-0,02	-0,03	0,95	0,054	2,1
<b>X<sub>4</sub> - Projetos de Pesquisa</b>	0,83	0,19	0,11	0,14	-0,06	-0,03	0,76	0,239	1,2
<b>X<sub>5</sub> - Linhas de Pesquisa</b>	0,91	0,16	0,12	0,14	-0,09	0,05	0,89	0,109	1,2
<b>X<sub>6</sub> - Publicações científicas</b>	0,77	0,08	0,23	0,42	-0,06	0,02	0,83	0,17	1,8
<b>X<sub>7</sub> - Pesquisas financiadas pelos setores produtivos</b>	0,29	-0,09	0,43	0,65	-0,15	0,01	0,72	0,284	2,4
<b>X<sub>8</sub> - Linhas de pesquisa com setores produtivos</b>	0,2	0,58	0,01	0,17	-0,35	0,43	0,72	0,282	3,1
<b>X<sub>9</sub> - Projetos de Pesquisas financiadas pelos setores produtivos</b>	0,32	0,12	0,25	0,84	-0,01	-0,05	0,89	0,112	1,5
<b>X<sub>10</sub> - Treinamento</b>	0,62	0,22	0,21	0,15	0,26	-0,07	0,57	0,427	2,1
<b>X<sub>11</sub> - Consultoria</b>	0,8	0,09	-0,06	-0,11	0,3	0,14	0,77	0,226	1,4
<b>X<sub>12</sub> - Projetos de extensão acadêmicos ou tecnológicos</b>	0,88	0,25	-0,01	-0,03	-0,08	0,07	0,84	0,16	1,2
<b>X<sub>13</sub> - Serviços de uso de laboratórios e equipamentos ; etc.</b>	0,92	0,12	0,07	0,08	0,06	-0,08	0,88	0,118	1,1

Continua...

Continua..

<b>Variáveis</b>	<b>RPC<sub>1</sub></b>	<b>RPC<sub>2</sub></b>	<b>RPC<sub>3</sub></b>	<b>RPC<sub>4</sub></b>	<b>RPC<sub>5</sub></b>	<b>RPC<sub>6</sub></b>	<b>h<sub>2</sub></b>	<b>u<sub>2</sub></b>	<b>com</b>
<b>X<sub>14</sub> - Treinamentos demandados pelos setores produtivos</b>	-0,08	-0,04	-0,08	-0,07	0,06	-0,66	0,47	0,535	1,1
<b>X<sub>15</sub> - Consultorias demandadas pelos setores produtivos</b>	0,12	-0,07	0,03	0,11	0,9	0,11	0,85	0,155	1,1
<b>X<sub>16</sub> - Projetos de extensão demandados pelos setores produtivos</b>	0,07	-0,07	-0,17	0,85	0,24	0,1	0,82	0,179	1,3
<b>X<sub>17</sub> - Serviços de uso de laboratórios e equipamentos ; etc. demandados pelos setores produtivos</b>	-0,09	-0,04	0,04	-0,08	0,24	0,72	0,59	0,411	1,3
<b>X<sub>18</sub> - Marcas (concedidas)</b>	-0,01	0,08	0,88	-0,02	-0,05	0	0,78	0,217	1
<b>X<sub>19</sub> - Desenhos Industriais (concedidos)</b>	0,24	0,9	-0,03	-0,03	0,16	-0,02	0,9	0,103	1,2
<b>X<sub>20</sub> - Patentes (publicadas)</b>	0,25	0,35	0,65	0,41	0,12	0,02	0,8	0,2	2,8
<b>X<sub>21</sub> - Programas de computador (concedidos)</b>	0,2	-0,06	0,73	0,04	-0,09	0,08	0,6	0,404	1,2
<b>X<sub>22</sub> - Desenhos Industriais (concedidos) demandados pelos setores produtivos</b>	0,17	0,94	-0,11	-0,02	-0,03	0,01	0,93	0,073	1,1

Continua...

Continua...

Variáveis	RPC <sub>1</sub>	RPC <sub>2</sub>	RPC <sub>3</sub>	RPC <sub>4</sub>	RPC <sub>5</sub>	RPC <sub>6</sub>	h <sub>2</sub>	u <sub>2</sub>	com
X <sub>23</sub> - Patentes (publicadas) demandadas pelos setores produtivos	0,22	0,55	0,53	0	-0,17	-0,03	0,66	0,344	2,5
X <sub>24</sub> - Programas de computador (concedidos) demandados pelos setores produtivos	0,11	0,88	0,26	-0,01	-0,07	-0,03	0,86	0,144	1,2
<b>SS loadings</b>	6,68	3,48	3,24	2,55	1,45	1,24			
<b>Proportion Var</b>	0,28	0,15	0,13	0,11	0,06	0,05			
<b>Cumulative Var</b>	0,28	0,42	0,56	0,66	0,73	0,78			
<b>Proportion Explained</b>	0,36	0,19	0,17	0,14	0,08	0,07			
<b>Cumulative Proportion</b>	0,36	0,55	0,72	0,86	0,93	1			

Nota: Média de complexidade do item = 1,6; Teste da hipótese de que 6 componentes são suficientes foi validada. A raiz quadrada média dos resíduos (RMSR) é 0,05 com o qui-quadrado empírico 67,28 com probabilidade < 1. Ajuste baseado em valores fora da diagonal = 0,98. A análise de componentes principais (PCA - Principal Components Analysis) é uma técnica estatística utilizada para reduzir a dimensionalidade dos dados, transformando um conjunto de possíveis variáveis correlacionadas em um conjunto menor de variáveis não correlacionadas chamadas componentes principais. No output que você forneceu, há informações relevantes sobre os componentes extraídos e as cargas de cada variável nesses componentes. Vamos decifrar as partes mais importantes do seu output: Standardized loadings (pattern matrix) based upon correlation matrix: Esta é a matriz de cargas dos componentes principais. As cargas são coeficientes de correlação entre as variáveis originais (X<sub>1</sub> a X<sub>24</sub>) e os componentes. Os valores das cargas podem variar entre -1 e 1. Uma carga alta (absoluta) indica que a variável contribui significativamente para o componente em questão. **h<sub>2</sub>**: representa a comunalidade da variável. Seu valor representa a proporção da variância em uma determinada variável que pode ser explicada pelos componentes. Um valor de **h<sub>2</sub>** próximo de 1 sugere que quase toda a variância da variável é contabilizada pelos componentes, enquanto um valor próximo de 0 sugere o contrário. **u<sub>2</sub>**: representa a unicidade da variável. Seu resultado é basicamente 1 - **h<sub>2</sub>**. Representa a proporção da variância da variável que NÃO é explicada pelos componentes. **com**: é a medida de complexidade da variável. Seu resultado indica quantos componentes são necessários para explicar a variável. Um valor próximo de 1 indica que a variável é basicamente uma função de um único componente, enquanto um valor maior sugere que a variável é uma combinação complexa de vários componentes. **SS loadings**: São as somas dos quadrados das cargas para cada componente. **Proportion Var**: Proporção da variância total explicada por cada componente. **Cumulative Var**: Proporção acumulada da variância explicada até o componente atual. **Proportion Explained**: Proporção da variância total explicada por cada

componente em relação à variância total que pode ser explicada por todos os componentes. **Cumulative Proportion**: Proporção acumulada da variância explicada até o componente atual, em relação à variância total que pode ser explicada por todos os componentes. **RMSR**: *Root Mean Square of Residuals*. Uma medida de ajuste do modelo. Valores baixos de **RMSR** indicam bom ajuste. O output que você forneceu mostra os resultados da análise de componentes principais para 24 variáveis, onde 6 componentes foram extraídos e rotacionados usando o método de **varimax**. Cada componente é uma combinação linear das variáveis originais, e o objetivo da rotação é tornar a interpretação dos componentes mais fácil, maximizando as altas cargas em algumas variáveis e minimizando as outras.

Tabela B. 7: Resultado da *Data Mining* Integrada de Canais de PD&S em ICTs 2008 a 2012 (Cenário 12, Abordagem 1)

		<i>Data Mining</i>
Fator 6	Resultados das análises HC, NHC e PCA	ICTs-SPP
Fator 5	Atividades Diversificadas para SPP	ICTs-SPP
Fator 4	Consultoria e Treinamento para SPP	ICTs-SPP
Fator 3	Projetos Colaborativos Financiados pelos SPP	GOV
Fator 2	Iniciativas de Propriedade Intelectual e captação de Investimento	ICTs
Fator 1	Cooperação em Inovação com SPP	ICTs-SPP
Fator 1	Pesquisa e Desenvolvimento Tradicionais	ICTs
	Informações Geográficas	GOV

Continua...

IIP_PD& SOrder	5	26	38	40
K_NHC	1	3	3	3
K_HC	1	2	2	2
X17	0	0	0	0
X15	0	0	0	0
X14	0	0	0	0
X7	427	106	189	17
X9	57	10	5	3
X16	0	2	0	0
X1	54328702	63339435	0	1259585
X21	38	0	0	0
X20	35	2	0	1
X18	31	0	0	0
X8	0	5	1	0
X24	1	0	0	0
X23	1	0	0	0
X22	0	0	0	0
X19	0	0	0	0
X13	156	54	54	32
X12	119	121	17	30
X11	61	44	40	25
X10	24	13	8	11
X6	2865	2409	1278	357
X5	480	313	278	162
X4	2998	1796	1025	957
X3	10774	4955	2068	887
X2	929	728	740	322
W3	CEFFET/M	CEFFET/RJ	CPII	IFAC
W2	MG	RJ	RJ	AC
W1	Sudeste	Sudeste	Sudeste	Norte

Continua...

Continua...

IIP_PD& S Order	34	29	39	35	11
K_NHC	3	3	3	3	2
K_HC	2	2	2	2	2
X17	0	0	0	0	0
X15	0	0	0	0	2
X14	1	0	0	0	0
X7	22	224	10	54	137
X9	21	9	6	7	29
X16	0	0	0	0	0
X1	17354220	37804535	13676564	69310682	58800734
X21	5	0	0	3	13
X20	0	6	0	0	26
X18	0	0	0	0	4
X8	0	0	0	0	0
X24	0	0	0	0	0
X23	0	0	0	0	0
X22	0	0	0	0	0
X19	0	0	0	0	1
X13	128	120	36	58	173
X12	65	52	23	58	162
X11	45	51	54	57	96
X10	17	19	18	17	24
X6	1323	1147	207	858	2024
X5	361	234	71	321	710
X4	2964	2574	805	1403	3627
X3	4096	4344	783	1974	6850
X2	844	818	287	611	1257
W3	IFAL	IFAM	IFAP	IFB	IFBA
W2	AL	AM	AP	DF	BA
W1	Nordeste	Norte	Norte	Centro	Nordeste

Continua...

Continua...



IIP_PD& S_Order	31	14	3	2	28
K_NHC	3	2	1	1	3
K_HC	2	2	1	1	2
X17	0	0	0	0	0
X15	0	0	1	0	0
X14	1	0	0	0	0
X7	128	12	501	209	9
X9	20	17	206	122	26
X16	0	0	5	0	1
X1	37299960	28098403	15519457	56194080	53608228
X21	1	2	5	14	4
X20	10	2	28	50	0
X18	0	0	0	9	0
X8	0	1	1	4	1
X24	0	0	0	0	0
X23	0	0	0	1	0
X22	0	0	0	0	0
X19	0	0	0	2	0
X13	191	263	266	228	56
X12	83	205	218	200	65
X11	86	189	134	137	75
X10	17	16	33	48	14
X6	1516	2193	4181	3927	1821
X5	299	583	685	820	304
X4	2130	4036	5336	4883	2236
X3	3396	7636	13108	11106	3340
X2	802	1014	1575	1462	852
W3	IFBAIAN	IFC	IFCE	IFES	IPF
W2	BA	SC	CE	ES	RJ
W1	Nordeste	Sul	Nordeste	Sudeste	Sudeste

Continua...

Continua...

IIP_PD&S Order	13	16	20	12	19
K_NHC	2	2	2	1	2
K_HC	2	2	2	2	2
X17	0	0	0	0	0
X15	0	0	1	0	0
X14	0	0	0	0	0
X7	6	35	20	344	29
X9	4	35	20	19	15
X16	0	0	0	0	0
X1	35729792	78442764	12576728	17147485	15958681
X21	4	12	3	0	23
X20	2	6	1	7	4
X18	7	4	0	0	3
X8	0	2	0	0	0
X24	0	0	0	0	0
X23	0	0	0	0	0
X22	0	0	0	0	0
X19	2	0	0	0	0
X13	261	162	167	203	151
X12	138	136	114	129	176
X11	217	99	76	99	143
X10	28	20	7	25	32
X6	1897	2183	4148	2230	1522
X5	389	595	496	623	360
X4	4313	3382	4004	6797	2709
X3	6254	6402	7974	8268	6116
X2	893	1152	787	1537	938
W3	IFFAR	IFGO	IFGOIAN	IFMA	IFMG
W2	RS	GO	GO	MA	MG
W1	Sul	Centro	Centro	Nordeste	Sudeste

Continua...

Continua...

IIP_PD&S Order	37	21	32	10	8
K_NHC	3	2	3	2	2
K_HC	2	2	2	2	2
X17	0	0	0	1	0
X15	0	0	0	2	3
X14	0	0	0	0	0
X7	12	41	62	148	55
X9	5	16	14	15	14
X16	0	1	0	1	1
X1	24963903	89735992	24457285	36426122	81601753
X21	2	4	0	0	18
X20	1	3	0	3	19
X18	0	0	1	1	2
X8	1	0	0	0	0
X24	0	0	0	0	0
X23	0	0	0	0	0
X22	0	0	0	0	0
X19	0	0	0	0	0
X13	68	123	80	203	137
X12	53	73	95	168	185
X11	57	111	67	160	88
X10	7	32	15	23	11
X6	803	2191	1145	1823	2359
X5	244	344	300	466	474
X4	2066	3963	2712	3681	4187
X3	2487	6658	3548	4875	7771
X2	558	1000	704	1183	1181
W3	IFMS	IFMT	IFNMG	IFPA	IFPB
W2	MS	MT	MG	PA	PB
W1	Centro	Centro	Sudeste	Norte	Nordeste

Continua...

Continua...

IIP_PD&S Order	9	22	18	17	7
K_NHC	2	2	2	2	1
K_HC	2	2	2	2	1
X17	0	0	0	0	0
X15	4	1	0	0	0
X14	0	0	0	0	0
X7	76	32	14	281	342
X9	18	8	9	63	61
X16	1	0	1	0	0
X1	11705622	14465097	65773660	50145371	14485565
X21	4	4	0	1	56
X20	4	9	0	3	10
X18	0	0	0	0	0
X8	0	0	0	1	3
X24	0	0	0	0	0
X23	0	0	0	0	0
X22	0	0	0	0	0
X19	2	0	0	0	0
X13	182	69	227	176	187
X12	154	118	182	187	193
X11	154	82	106	55	103
X10	55	14	13	17	41
X6	1684	1734	2618	2687	3088
X5	429	345	646	679	674
X4	2333	3394	4930	3773	6261
X3	5589	6025	6339	7781	11687
X2	1016	1138	1250	1016	1527
W3	IPPE	IPPI	IPPR	IFRJ	IFRN
W2	PE	PI	PR	RJ	RN
W1	Nordeste	Nordeste	Sul	Sudeste	Nordeste

Continua...

Continua...

IIP_PD&S Order	27	41	6	1	24
K_NHC	3	3	1	2	3
K_HC	2	2	1	1	2
X17	0	0	0	0	1
X15	0	0	0	0	0
X14	0	0	0	0	0
X7	15	20	101	65	60
X9	5	1	18	49	15
X16	0	0	0	0	0
X1	37458420	48510942	12005317	83642391	12950852
X21	2	0	7	4	12
X20	9	0	1	25	12
X18	0	0	0	0	5
X8	2	1	3	6	4
X24	0	0	0	2	0
X23	0	0	1	1	0
X22	0	0	0	1	0
X19	0	0	0	7	0
X13	79	29	335	271	60
X12	67	5	493	451	30
X11	90	15	149	150	73
X10	21	2	36	35	13
X6	1190	291	2676	3096	919
X5	310	50	808	905	267
X4	3308	958	7367	6837	1828
X3	2926	1478	9049	6380	2455
X2	670	269	1333	1552	519
W3	IFRO	IFRR	IFRS	IFSC	IFSE
W2	RO	RR	RS	SC	SE
W1	Norte	Norte	Sul	Sul	Nordeste

Continua...

Continua...

IIP_PD&S Order	23	25	4	36	15
K_NHC	2	3	1	3	2
K_HC	2	2	1	2	2
X17	0	0	0	0	0
X15	0	0	0	0	1
X14	0	0	0	0	0
X7	116	28	162	4	14
X9	14	15	20	5	20
X16	0	1	0	0	0
X1	51879076	12302270	13951395	13850544	60499279
X21	8	9	3	3	9
X20	6	4	11	8	10
X18	2	3	2	0	0
X8	1	0	1	0	2
X24	0	0	0	0	0
X23	0	1	0	0	0
X22	0	0	0	0	0
X19	2	0	0	0	0
X13	139	123	331	71	115
X12	99	90	659	50	112
X11	80	53	192	37	130
X10	20	20	31	13	28
X6	1448	1506	4180	967	1798
X5	265	271	1318	209	342
X4	3231	2603	4533	1934	3499
X3	5895	4830	11048	2277	7004
X2	647	607	2298	457	945
W3	IFSEMG	IFSMG	IFSP	IFSTPE	IFSUL
W2	MG	MG	SP	PE	RS
W1	Sudeste	Sudeste	Sudeste	Nordeste	Sul

Continua...

Continua...

IIP_PD&S Order	30	33
K_NHC	2	3
K_HC	2	2
X17	0	0
X15	0	0
X14	1	0
X7	8	45
X9	22	5
X16	0	0
X1	26547882	41942126
X21	1	0
X20	1	0
X18	0	0
X8	0	0
X24	0	0
X23	0	0
X22	0	0
X19	0	0
X13	130	84
X12	114	66
X11	78	111
X10	25	13
X6	1545	1161
X5	372	264
X4	4122	2239
X3	4961	3078
X2	644	642
W3	IFTM	IFTO
W2	MG	TO
W1	Sudeste	Norte

Continua...

Quadro B. 16: Lista dos códigos ou atividades econômicas na CNAE

<b>Seção</b>	<b>Divisões</b>	<b>Denominação</b>
<b>A</b>	01 .. 03	agricultura, pecuária, produção florestal, pesca e aquicultura
<b>B</b>	05 .. 09	indústrias extrativas
<b>C</b>	10 .. 33	indústrias de transformação
<b>D</b>	35 .. 35	eletricidade e gás
<b>E</b>	36 .. 39	água, esgoto, atividades de gestão de resíduos e descontaminação
<b>F</b>	41 .. 43	construção
<b>G</b>	45 .. 47	comércio; reparação de veículos automotores e motocicletas
<b>H</b>	49 .. 53	transporte, armazenagem e correio
<b>I</b>	55 .. 56	alojamento e alimentação
<b>J</b>	58 .. 63	informação e comunicação
<b>K</b>	64 .. 66	atividades financeiras, de seguros e serviços relacionados
<b>L</b>	68 .. 68	atividades imobiliárias
<b>M</b>	69 .. 75	atividades profissionais, científicas e técnicas
<b>N</b>	77 .. 82	atividades administrativas e serviços complementares
<b>O</b>	84 .. 84	administração pública, defesa e seguridade social
<b>P</b>	85 .. 85	educação
<b>Q</b>	86 .. 88	saúde humana e serviços sociais
<b>R</b>	90 .. 93	artes, cultura, esporte e recreação
<b>S</b>	94 .. 96	outras atividades de serviços
<b>T</b>	97 .. 97	serviços domésticos
<b>U</b>	99 .. 99	organismos internacionais e outras instituições extraterritoriais



Quadro B. 17: Comparação de Funcionalidades e Disponibilidade de Informações: Análise do Dashboard Proposto versus Portal Integra, Portal Universidade 360 e StelaExperta®

<b>Id</b>	<b>Funcionalidades</b>	<b>PD&amp;S Analytics</b>	<b>Portal Integra</b>	<b>Portal Univ. 360</b>	<b>StelaExperta®</b>
1	Acompanhamento e monitoramento IIP_PD&S	Sim	Não	Não	Não
2	Adaptável às necessidades dos tomadores de decisão	Sim	Sim	Não	Sim
3	Análise baseada em custos, tempo e qualidade	Sim	Sim	Sim	Sim
4	Análise de métricas simultâneas	Sim	Não	Não	Não
5	Análise por Setor Tripla Hélice	Sim	Não	Não	Não
6	Atualizações e adaptações contínuas	Sim	Sim	Sim	Sim
7	Busca textual dinâmica	Sim	Não	Não	Sim
8	Cessão Infraestrutura	Sim	Sim	Não	Não
9	Comparativo interinstitucional	Sim	Não	Não	Não
10	Consultorias	Sim	Não	Não	Não
11	Dados relacionados a bolsas e financiamentos	Sim	Não	Não	Sim
12	Desenho Industrial	Sim	Sim	Não	Não
13	Desenvolvimento tecnológico	Sim	Sim	Sim	Sim
14	Exportação de dados e relatórios	Sim	Não	Não	Sim
15	Extensão Acadêmica	Sim	Não	Não	Não
16	Filtros adicionais	Sim	Não	Não	Sim
17	Filtros básicos	Sim	Sim	Sim	Sim
18	Filtros específicos por métrica	Sim	Não	Não	Não
19	Fomento Federal em P&D	Sim	Não	Não	Não
20	Grupos de pesquisa	Sim	Sim	Sim	Sim
21	Indicadores de relevância sugeridos pelo usuário	Sim	Não	Não	Não
22	Integração com dados institucionais de PD&S	Sim	Não	Não	Não
23	Linhas de pesquisa	Sim	Sim	Sim	Sim

Continua...

Continua...

<b>Id</b>	<b>Funcionalidades</b>	<b>PD&amp;S Analytics</b>	<b>Portal Integra</b>	<b>Portal Univ. 360</b>	<b>StelaExperta®</b>
24	Mapa de densidade	Sim	Não	Não	Não
25	Marca	Sim	Não	Não	Não
26	Módulo de análise comparativo	Sim	Não	Não	Não
27	Módulo de análise detalhada	Sim	Não	Não	Não
28	Padronização de métricas e KPIs	Sim	Não	Não	Não
29	Padronização de valores com z-score	Sim	Não	Não	Não
30	Patente	Sim	Sim	Sim	Sim
31	Perguntas estratégicas inteligente	Sim	Não	Não	Não
32	Pesquisador	Sim	Sim	Não	Sim
33	Pesquisas com potencial de Des. Tecn. (DT)	Sim	Não	Não	Não
34	Programa Computador	Sim	Sim	Não	Sim
35	Projetos de pesquisa	Sim	Não	Sim	Sim
36	Publicações Científicas	Sim	Não	Sim	Sim
37	Tabela de fatos interativa	Sim	Não	Não	Não
38	Treinamentos	Sim	Não	Não	Não
39	Visão geral das métricas e KPIs	Sim	Sim	Sim	Sim
40	Visualização de métricas por grande área do conhecimento	Sim	Não	Sim	Sim
41	Visualização de métricas por set. econômicos (CNAE)	Sim	Não	Não	Não
42	Visualização interativa de gráficos, mapas e tabelas	Sim	Não	Não	Sim
43	Orientações de pesquisa	Sim	Não	Não	Sim
44	Vitrine tecnológica institucional	Não	Sim	Não	Não
45	Portfólio institucional	Não	Sim	Não	Não

Quadro B. 18: Estudos das variáveis contributivas para a separação dos agrupamentos HC, Cenário 12, Abordagem 1 e Abordagem 2 utilizando os testes de ANOVA, Benjamini Hochberg (BH), Bonferroni e Magnitude

Var	valor-p	Eta_squared	Bonferroni	BH	Res_ANOVA	Res_Magnitude	Res_Bonferroni	Res_BH
X <sub>1</sub>	0,05212283	0,093306457	1	0,062547395	no best fitted	Média	no best fitted	no best fitted
X <sub>2</sub>	4,30471E-06	0,422073471	0,000103313	2,06626E-05	best fitted	Grande	best fitted	best fitted
X <sub>3</sub>	2,3006E-07	0,500701766	5,52144E-06	4,45571E-06	best fitted	Grande	best fitted	best fitted
X <sub>4</sub>	4,23646E-05	0,352829694	0,001016751	0,000101675	best fitted	Grande	best fitted	best fitted
X <sub>5</sub>	5,22882E-07	0,479750367	1,25492E-05	4,45571E-06	best fitted	Grande	best fitted	best fitted
X <sub>6</sub>	1,16399E-06	0,458526364	2,79357E-05	6,98393E-06	best fitted	Grande	best fitted	best fitted
X <sub>7</sub>	6,02185E-05	0,341537777	0,001445244	0,000131386	best fitted	Grande	best fitted	best fitted
X <sub>8</sub>	0,001461388	0,231207068	0,03507332	0,002505237	best fitted	Grande	best fitted	best fitted
X <sub>9</sub>	6,64551E-06	0,409472064	0,000159492	2,58741E-05	best fitted	Grande	best fitted	best fitted
X <sub>10</sub>	7,93246E-05	0,332567639	0,001903791	0,000158649	best fitted	Grande	best fitted	best fitted
X <sub>11</sub>	0,018867264	0,133402571	0,452814336	0,026636137	best fitted	Média	no best fitted	best fitted
X <sub>12</sub>	5,56964E-07	0,478104763	1,33671E-05	4,45571E-06	best fitted	Grande	best fitted	best fitted
X <sub>13</sub>	2,75075E-05	0,36645927	0,00066018	7,33533E-05	best fitted	Grande	best fitted	best fitted
X <sub>14</sub>	0,426996153	0,01625387	1	0,465813985	no best fitted	Média	no best fitted	no best fitted
X <sub>15</sub>	0,472408633	0,013320498	1	0,492948139	no best fitted	Média	no best fitted	no best fitted
X <sub>16</sub>	0,224560842	0,037582064	1	0,256640963	no best fitted	Média	no best fitted	no best fitted
X <sub>17</sub>	0,522637251	0,010558069	1	0,522637251	no best fitted	Média	no best fitted	no best fitted
X <sub>18</sub>	0,015173776	0,141964002	0,36417062	0,022760664	best fitted	Grande	no best fitted	best fitted
X <sub>19</sub>	0,031429404	0,113272166	0,754305703	0,0397003	best fitted	Média	no best fitted	best fitted
X <sub>20</sub>	7,54662E-06	0,405734019	0,000181119	2,58741E-05	best fitted	Grande	best fitted	best fitted
X <sub>21</sub>	0,001335074	0,234526831	0,032041775	0,002464752	best fitted	Grande	best fitted	best fitted
X <sub>22</sub>	0,025564238	0,121428571	0,613541702	0,03408565	best fitted	Média	no best fitted	best fitted
X <sub>23</sub>	1,34608E-05	0,388445378	0,000323059	4,03824E-05	best fitted	Grande	best fitted	best fitted
X <sub>24</sub>	0,001823545	0,22303207	0,043765091	0,002917673	best fitted	Grande	best fitted	best fitted

Nota: De acordo com Cohen (1988) a magnitude pode ser classificada como pequena ( $0,01 \leq \text{Eta\_squared} < 0,06$ ), média ( $0,06 \leq \text{Eta\_squared} < 0,14$ ) e grande ( $\text{Eta\_squared} \geq 0,14$ ).

Quadro B. 19: Síntese da Análise Hierárquica de Agrupamentos (HC) das ICTs por Indicadores para o Cenário 12, Abordagem 1

<i>k</i>	<i>n</i>	X <sub>1</sub> (R\$)	X <sub>2</sub>	X <sub>3</sub>	X <sub>4</sub>	X <sub>5</sub>	X <sub>6</sub>	X <sub>7</sub>	X <sub>8</sub>	X <sub>9</sub>	X <sub>10</sub>	X <sub>11</sub>	X <sub>12</sub>	X <sub>13</sub>	X <sub>14</sub>	X <sub>15</sub>	X <sub>16</sub>	X <sub>17</sub>	X <sub>18</sub>	X <sub>19</sub>	X <sub>20</sub>	X <sub>21</sub>	X <sub>22</sub>	X <sub>23</sub>	X <sub>24</sub>	...
1	7	23.184.643,12	1527	11048	5336	808	3096	209	3	57	35	137	218	266	0	0	0	0	0	0	25	7	0	1	0	...
2	33	10.208.072,96	818	4955	2964	342	1545	32	0	15	17	80	99	123	0	0	0	0	0	0	3	3	0	0	0	...

Continua...

<i>k</i>	...	W <sub>3</sub>	W <sub>2</sub>	W <sub>1</sub>	Média de <i>K</i>	Silhueta Média
1	...	CEFET/MG, IFCE, IFES, IFRN, IFRS, IFSC, IFSP	MG, CE, ES, RN, RS, SC, SP	Sudeste, Nordeste, Sul	-0,32	0,35
2	...	CEFET/RJ, IFAC, IFAL, IFAM, IFAP, IFB, IFBA, IFBAIANO, IFC, IFF, IFFAR, IFGO, IFGOIANO, IFMA, IFMG, IFMS, IFMT, IFNMG, IFPA, IFPB, IFPE, IFPI, IFPR, IFRJ, IFRO, IFRR, IFSE, IFSEMG, IFSMG, IFSTPE, IFSUL, IFTM, IFTO	RJ, AC, AL, AM, AP, DF, BA, SC, RS, GO, MA, MG, MS, MT, PA, PB, PE, PI, PR, RO, RR, SE, TO	Sudeste, Norte, Nordeste, Centro Oeste, Sul	0,49	

Quadro B. 20: Síntese da Análise Hierárquica de Agrupamentos (HC) das ICTs por Indicadores para o Cenário 12, Abordagem 2

<i>k</i>	<i>n</i>	X <sub>1</sub> (RS)	X <sub>2</sub>	X <sub>3</sub>	X <sub>4</sub>	X <sub>5</sub>	X <sub>6</sub>	X <sub>7</sub>	X <sub>8</sub>	X <sub>9</sub>	X <sub>10</sub>	X <sub>11</sub>	X <sub>12</sub>	X <sub>13</sub>	X <sub>14</sub>	X <sub>15</sub>	X <sub>16</sub>	X <sub>17</sub>	X <sub>18</sub>	X <sub>19</sub>	X <sub>20</sub>	X <sub>21</sub>
1	7	NA	1527	11048	5336	808	3096	209	3	57	35	137	218	266	0	0	25	7	0	1	0	1527
2	34	NA	810	4915	2838	331,5	1533,5	33,5	0	14,5	17	79	97	123	0	0	3	2,5	0	0	0	810

Continua...

X <sub>22</sub>	X <sub>23</sub>	X <sub>24</sub>	...
11048	5336	808	...
4915	2838	331,5	...

Continua...

<i>k</i>	...	W <sub>3</sub>	W <sub>2</sub>	W <sub>1</sub>	Média de <i>K</i>	Silhueta Média
1	...	CEFET/MG, IFCE, IFES, IFRN, IFRS, IFSC, IFSP	MG, CE, ES, RN, RS, SC, SP	Sudeste, Nordeste, Sul	0,46	0,48
2	...	CEFET/RJ, CPII, IFAC, IFAL, IFAM, IFAP, IFB, IFBA, IFBAIANO, IFC, IFF, IFFAR, IFGO, IFGOIANO, IFMA, IFMG, IFMS, IFMT, IFNMG, IFPA, IFPB, IFPE, IFPI, IFPR, IFRJ, IFRO, IFRR, IFSE, IFSEMG, IFSMG, IFSTPE, IFSUL, IFTM, IFTO	RJ, AC, AL, AM, AP, DF, BA, SC, RS, GO, MA, MG, MS, MT, PA, PB, PE, PI, PR, RO, RR, SE, TO	Sudeste, Norte, Nordeste, Centro Oeste, Sul	0,48	

Quadro B. 21: Síntese da Análise Não Hierárquica de Agrupamentos (NHC: *K-means*) das ICTs por Indicadores para o Cenário 12, Abordagem 1

<i>k</i>	<i>n</i>	X <sub>1</sub> (RS)	X <sub>2</sub>	X <sub>3</sub>	X <sub>4</sub>	X <sub>5</sub>	X <sub>6</sub>	X <sub>7</sub>	X <sub>8</sub>	X <sub>9</sub>	X <sub>10</sub>	X <sub>11</sub>	X <sub>12</sub>	X <sub>13</sub>	X <sub>14</sub>	X <sub>15</sub>	X <sub>16</sub>	X <sub>17</sub>	X <sub>18</sub>	X <sub>19</sub>	X <sub>20</sub>	X <sub>21</sub>	X <sub>22</sub>	X <sub>23</sub>	X <sub>24</sub>	...
1	4	1182	9272,5	3934	477	2612	178,5	0	36	23,5	112,5	176,5	179,5	0	1	0,5	0	5,5	0	27	16	0	0,5	0	1182	...
2	37	844	5589	3231	344	1684	41	0	15	18	82	112	128	0	0	0	0	0	0	3	3	0	0	0	844	...

Continua...

<i>k</i>	...	W <sub>3</sub>	W <sub>2</sub>	W <sub>1</sub>	Média de <i>K</i>	Silhueta Média
1	...	CEFET/MG, IFES, IFPA, IFPB	MG, ES, PA, PB	Sudeste, Norte, Nordeste	0,49	0,84
2	...	CEFET/RJ, CPII, IFAC, IFAL, IFAM, IFAP, IFB, IFBA, IFBAIANO, IFC, IFCE, IFF, IFFAR, IFGO, IFGOIANO, IFMA, IFMG, IFMS, IFMT, IFNMG, IFPE, IFPI, IFPR, IFRJ, IFRN, IFRO, IFRR, IFRS, IFSC, IFSE, IFSEMG, IFSMG, IFSP, IFSTPE, IFSUL, IFTM, IFTO	RJ, AC, AL, AM, AP, DF, BA, SC, CE, RS, GO, MA, MG, MS, MT, PE, PI, PR, RN, RO, RR, SE, SP, TO	Sudeste, Norte, Nordeste, Centro Oeste, Sul	0,88	

Quadro B. 22 Síntese da Análise Não Hierárquica de Agrupamentos (NHC: Kmeans) das ICTs por Indicadores para o Cenário 12, Abordagem 2

<i>k</i>	<i>n</i>	X <sub>1</sub> (RS)	X <sub>2</sub>	X <sub>3</sub>	X <sub>4</sub>	X <sub>5</sub>	X <sub>6</sub>	X <sub>7</sub>	X <sub>8</sub>	X <sub>9</sub>	X <sub>10</sub>	X <sub>11</sub>	X <sub>12</sub>	X <sub>13</sub>	X <sub>14</sub>	X <sub>15</sub>	X <sub>16</sub>	X <sub>17</sub>	X <sub>18</sub>	X <sub>19</sub>	X <sub>20</sub>	X <sub>21</sub>	X <sub>22</sub>	X <sub>23</sub>	X <sub>24</sub>	...
1	7	NULL	1527	11048	5336	685	3088	342	1	57	33	134	200	228	0	0	11	7	0	0	0	7		1527	11048	...
2	17	NULL	642	2926	2066	271	1147	28	0	7	14	54	58	68	0	0	1	1	0	0	0	17		642	2926	...
3	17	NULL	1016	6380	3773	466	2024	35	0	17	23	106	154	167	0	0	4	4	0	0	0	17		1016	6380	...

Continua...

<i>k</i>	...	W <sub>3</sub>	W <sub>2</sub>	W <sub>1</sub>	Média de <i>K</i>	Silhueta Média	Índice Comparativo com Quadro 4
1	...	CEFET/MG, IFCE, IFES, IFMA, IFRN, IFRS, IFSP	MG, CE, ES, MA, RN, RS, SP	Sudeste, Nordeste, Sul	0,49	0,48	0,27
2	...	CEFET/RJ, CPII, IFAC, IFAL, IFAM, IFAP, IFB, IFBAIANO, IFF, IFMS, IFNMG, IFRO, IFRR, IFSE, IFSMG, IFSTPE, IFTO	RJ, AC, AL, AM, AP, DF, BA, MS, MG, RO, RR, SE, PE, TO	Sudeste, Norte, Nordeste, Centro Oeste	0,35		
3	...	IFBA, IFC, IFFAR, IFGO, IFGOIANO, IFMG, IFMT, IFPA, IFPB, IFPE, IFPI, IFPR, IFRJ, IFSC, IFSEMG, IFSUL, IFTM	BA, SC, RS, GO, MG, MT, PA, PB, PE, PI, PR, RJ	Nordeste, Sul, Centro Oeste, Sudeste, Norte	0,51		

Tabela B. 8: Análise Comparativa da Evolução e Posicionamento do Brasil em Indicadores de Pesquisa e Inovação: 2008-2018

Período	Index	N	in	out	Evolução	Brasil	MUND	TOP5 <sup>(1)</sup>	ALC	BRICS
#	Descrição	Ranking (Brasil)			%YOY					
2008 to 2017	<i>Quality Of Scientific Research Institutions</i>	151	43	77	-34	-1,74	0,62	-0,11	-0,19	0,39
2008 to 2018	<i>R&amp;D Transfer</i>	112	NA	NA	NA	1,36	0,41	0,3	0,73	1,42
2013 to 2018	<i>Human capital and research</i>	213	75	52	23	3,01	3,68	0,7	1,44	4,34
2008 to 2018	<i>Research And Development Expenditure</i>	149	NA	NA	NA	0,29	2,59	0,45	3,34	1,33
2013 to 2018	<i>Gross Expenditure On R&amp;D (GERD)</i>	142	31	27	4	1,67	-0,21	0,33	1,95	0,9
2013 to 2018	<i>High-Tech Exports</i>	159	44	35	9	6,83	-1,98	-7,06	-3,77	1,88
2013-2018	<i>Global Innovation Index <sup>(1)</sup></i>	215	64	64	0	-1,62	1,91	0,27	-0,32	0,8
2013-2018	<i>High-Tech And Medium High-Tech Output</i>	119	22	30	-8	-4,84	-2,18	-1,83	-4,13	-2,84
2013-2018	<i>Research And Development (R&amp;D)</i>	211	33	28	5	7,96	6,33	2,77	6,24	5,54
2013-2018	<i>University/industry research collaboration</i>	213	42	67	-25	-4,41	0,95	-0,31	-2,42	0,65
2008 to 2017	<i>Capacity For Innovation</i>	151	27	73	-46	0,22	2,69	0,76	3,45	1,85
2008 to 2017	<i>University-Industry Collaboration In Research &amp; Development</i>	151	50	70	-20	-0,15	0,6	0,16	0,93	0,69
2013-2018	<i>Scientific and technical publications</i>	125	57	54	3	4,26	3,14	-0,03	8	5,33

Nota: O **Index** refere-se a um indicador ou métrica. **N** refere-se ao número total de países ou entidades analisados dentro de um determinado índice. **In** representa o número de países ou entidades que entraram ou melhoraram sua posição em um ranking específico durante um período determinado. **Out** representa o número de países ou entidades que saíram ou pioraram sua posição em um ranking específico durante um período determinado. **Evolução** refere-se à mudança líquida ou diferença entre as entidades que entraram ("in") e as que saíram ("out"). É a representação do progresso ou retrocesso em um índice específico. **MUND** refere-se ao desempenho ou mudança percentual médio mundial em relação a um índice específico durante um período determinado. **TOP5** representa o desempenho ou mudança percentual médio dos cinco principais países ou entidades em um índice específico (Suíça, Holanda, Suécia, Reino Unido, Estados Unidos da América). **ALC** representa o desempenho ou mudança percentual médio dos países da América Latina e do Caribe em um índice específico. **BRICS** refere-se ao desempenho ou mudança percentual médio dos países do grupo BRICS (Brasil, Rússia, Índia, China e África do Sul) em relação a um índice específico.



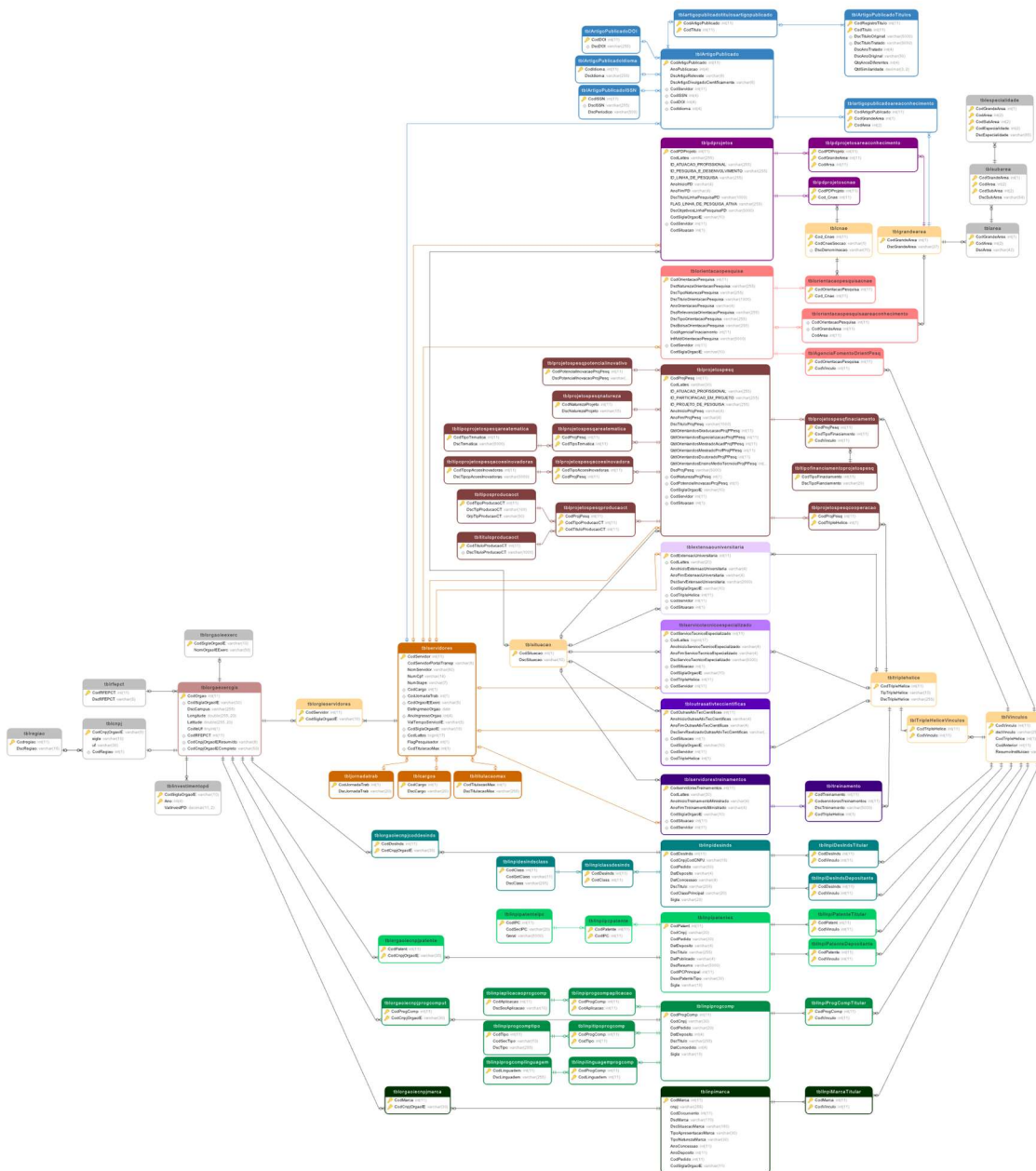


Figura B. 1: Representação visual do modelo de entidade de relacionamento (MER) utilizado para criar o Data Warehouse (DW) do projeto.

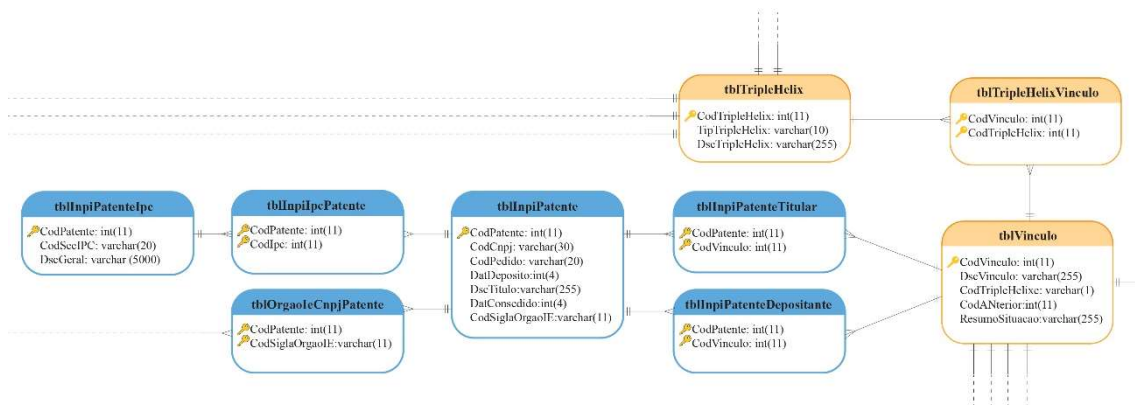


Figura B. 2: Esquema de uma pequena fração do banco de dados temporário do modelo de entidade de relacionamento utilizada para a carga do Data Warehouse

Nota: A Figura acima, apresenta-se um fragmento do Modelo de Entidade-Relacionamento após a Subseção 3.3. Nessa ilustração, as tabelas que contêm dados sobre patentes estão marcadas em azul, enquanto as tabelas intermediárias, que estabelecem conexões entre os dados de patentes e os atores do modelo de TH, através dos vínculos formados pelos titulares e/ou depositantes com as Instituições de ICTs, são destacadas em laranja claro. Existem, ainda, outras relações entre as tabelas, indicadas por linhas tracejadas. Contudo, devido à limitação de espaço, tais relações foram omitidas para facilitar a visualização completa do MER.

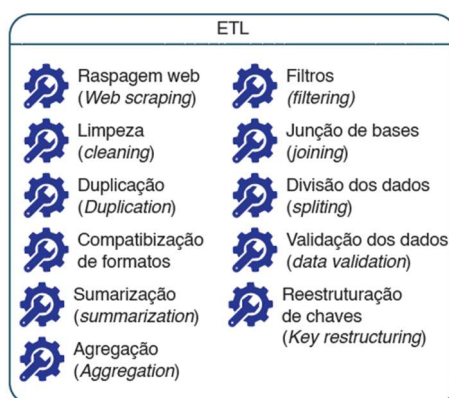


Figura B. 3: Apresentação do procedimento de ETL aplicado neste trabalho, adaptado de adaptado de Paikaray et al. (2021, p. 505-508)

Código B. 1: Fragmento do algoritmo criado no banco de dados temporário para carregamento no *Data Warehouse*. A carga corresponde aos dados de PD&S, entre 2008 e 2018, com 28 variáveis

```
CREATE TABLE `FatoPD&S` AS SELECT Invest.CodSiglaOrgaoIE AS Cod_instituicao, Invest.Ano AS Cod_tempo, Invest.ValInvestPD AS Val_totalValInvesPD, (SELECT count(*) from.mvp_tese.tblServidores AS Serv WHERE Serv.CodSiglaOrgaoIE = Invest.CodSiglaOrgaoIE ) AS Qtd_Pesquisadores, ... FROM.mvp_tese.tblInvestimentoPD AS Invest;
```

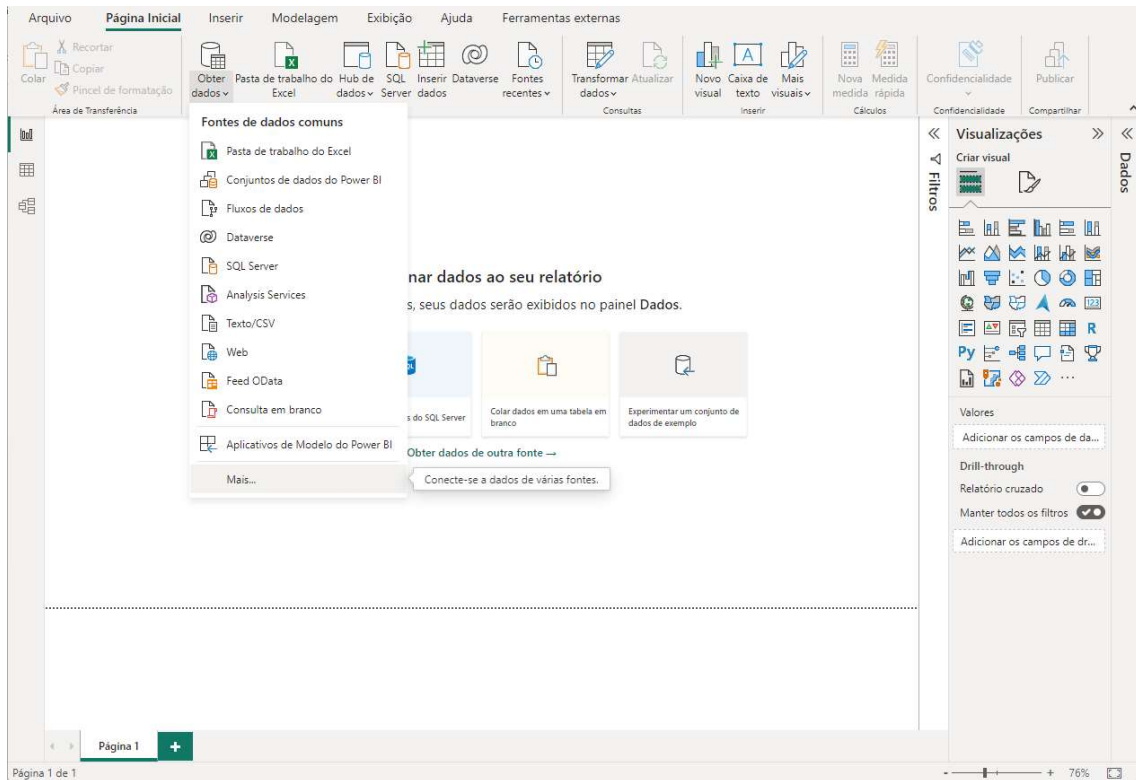


Figura B. 4: Esquema Geral para Carregar os dados no Power BI

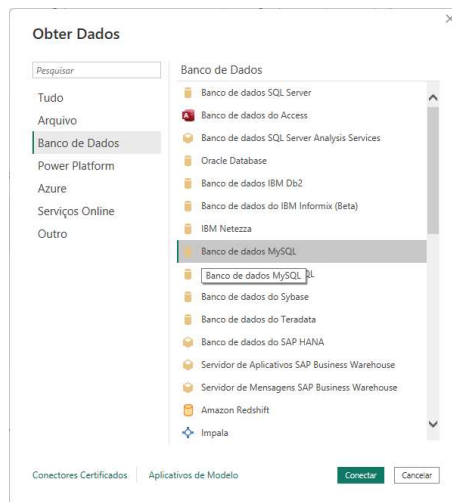


Figura B. 5: Seleção do tipo de banco de dado padrão MySQL

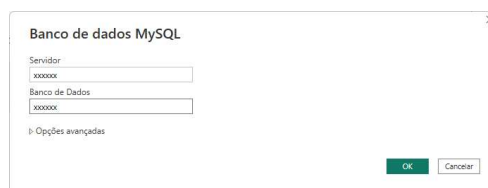


Figura B. 6: Acesso ao banco de dados do MySQL para carregar os dados no Power BI

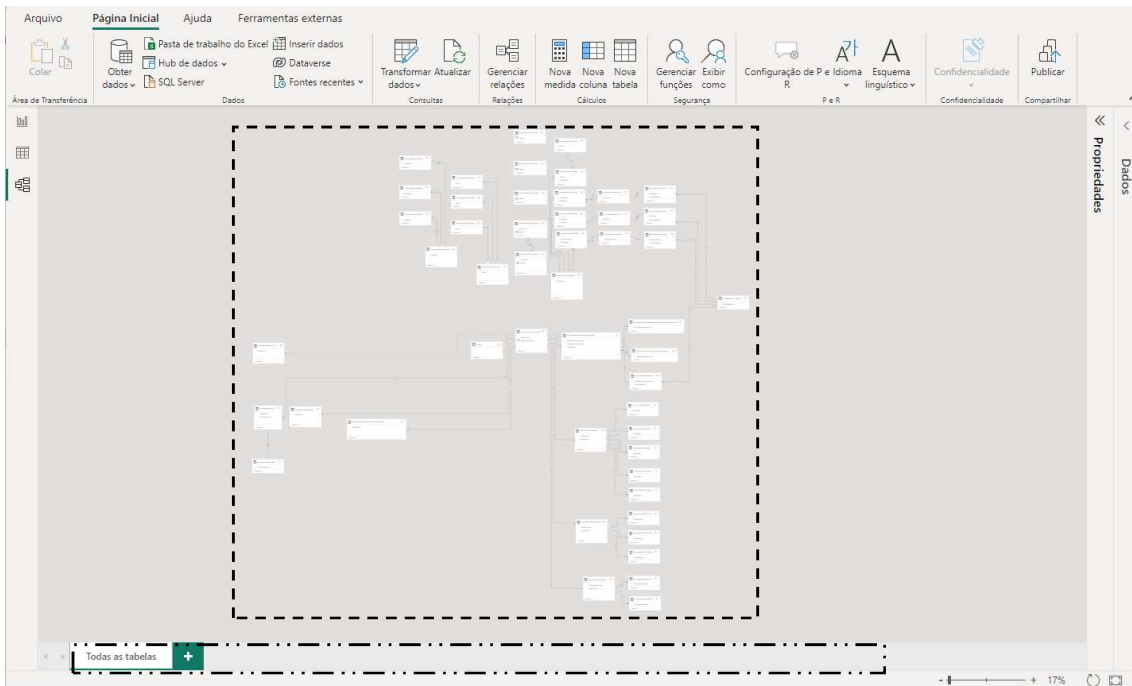


Figura B. 7: Tabelas de dados do MySQL carregadas no Power BI no início do projeto. Região tracejada destaca tabelas pré-tratadas no MySQL, enquanto região com dois pontos indica falta de modelagem dimensional no Data Warehouse do Power BI

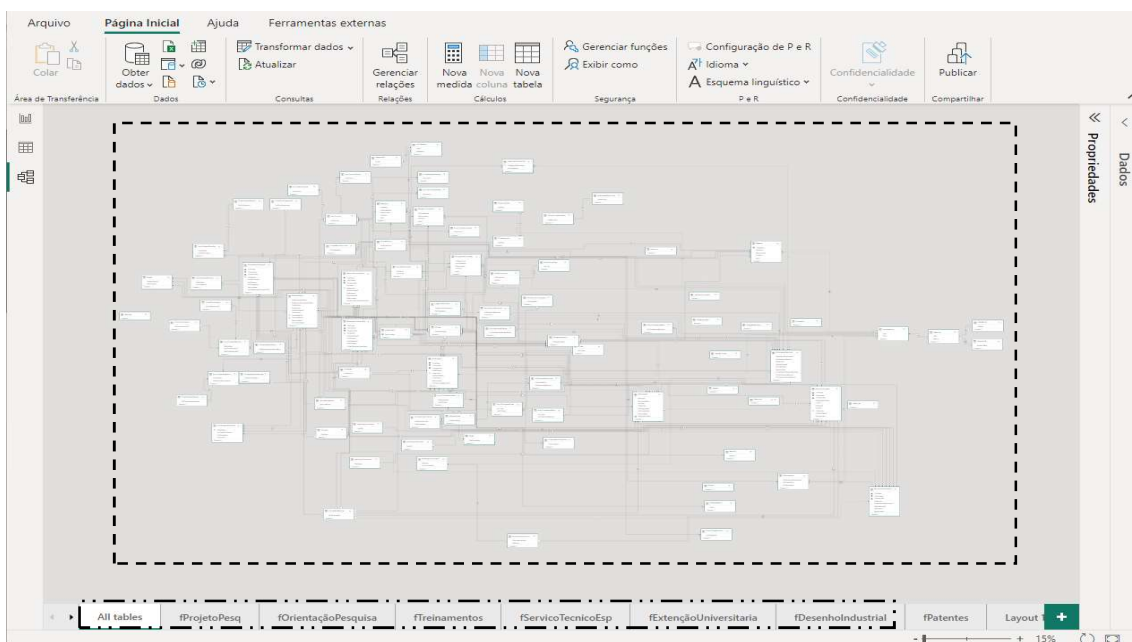


Figura B. 8: Ilustração das tabelas de dados pós-tratamento no Power BI (segundo ETL). Região tracejada mostra tabelas pós-tratamento. Região com dois pontos destaca várias modelagens dimensionais (tabelas de fatos e dimensões) implementadas no DW no final do projeto

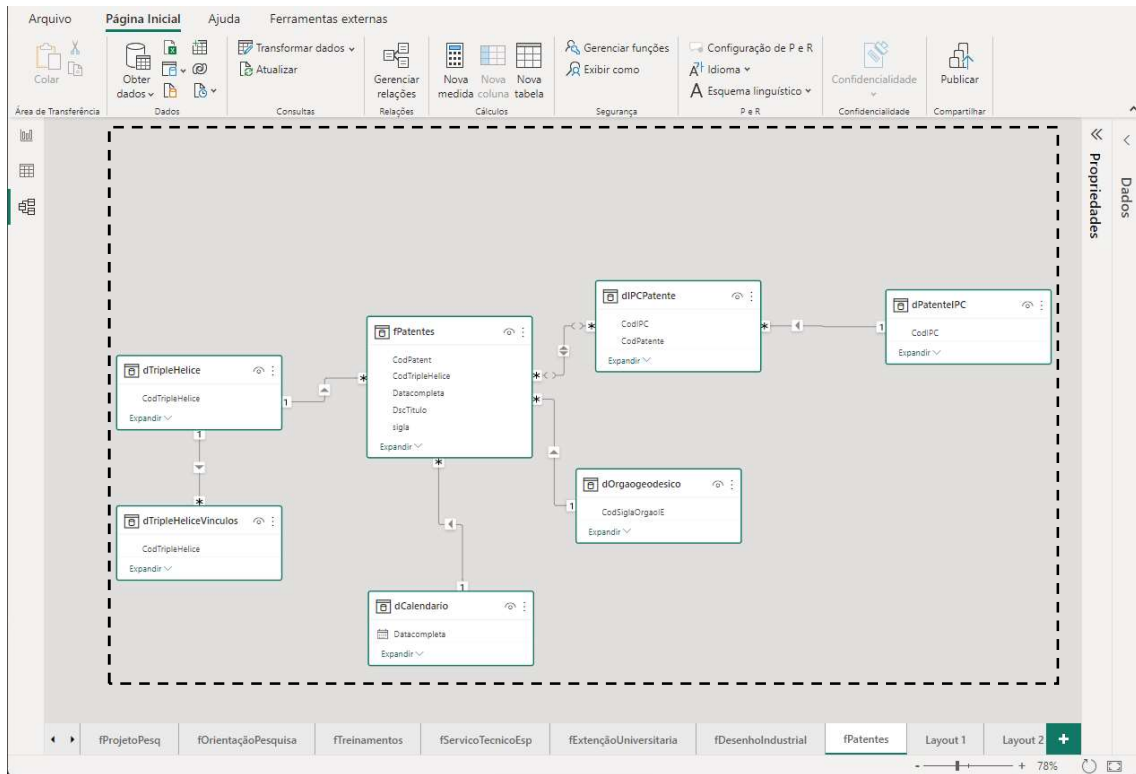


Figura B. 9: Ilustração da implementação da modelagem dimensional do esquema floco de neve da métrica Patente ( $X_{20}$  e  $X_{23}$ ) no Data Warehouse no Power BI. A região tracejada mostra as tabelas fato e dimensões organizadas. A região com dois pontos destaca a aba correspondente à modelagem dimensional no projeto

```

_metrica_orientacoes.pesquisa =
COUNTROWS(FILTER(fOrientacaoPesquisa,
fOrientacaoPesquisa[CodOrientacaoPesquisa]))

```

Figura B. 10: Ilustração da implementação da codificação de uma métrica no *Data Warehouse* do Power BI com DAX

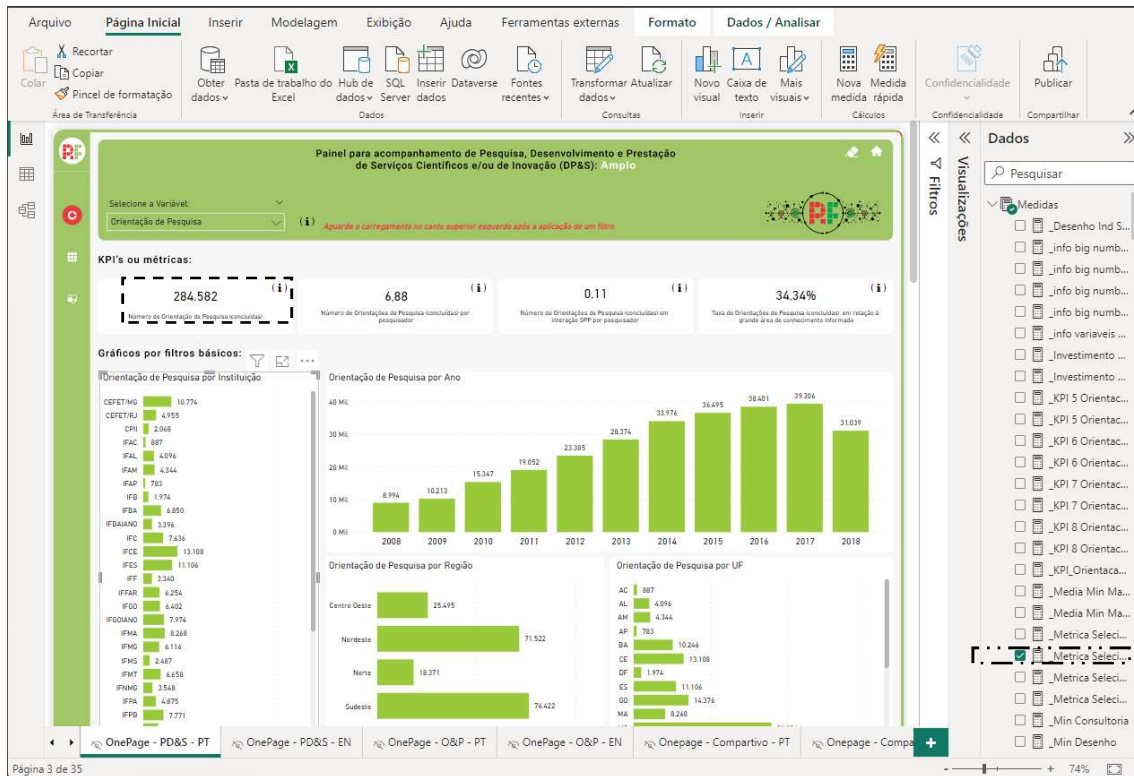


Figura B. 11: Ilustração da implementação da métrica "Total de Orientação de Pesquisa" utilizando DAX no *Data Warehouse* do Power BI. A região tracejada representa a visualização final e a com dois pontos indica a extraída da modelagem de dados

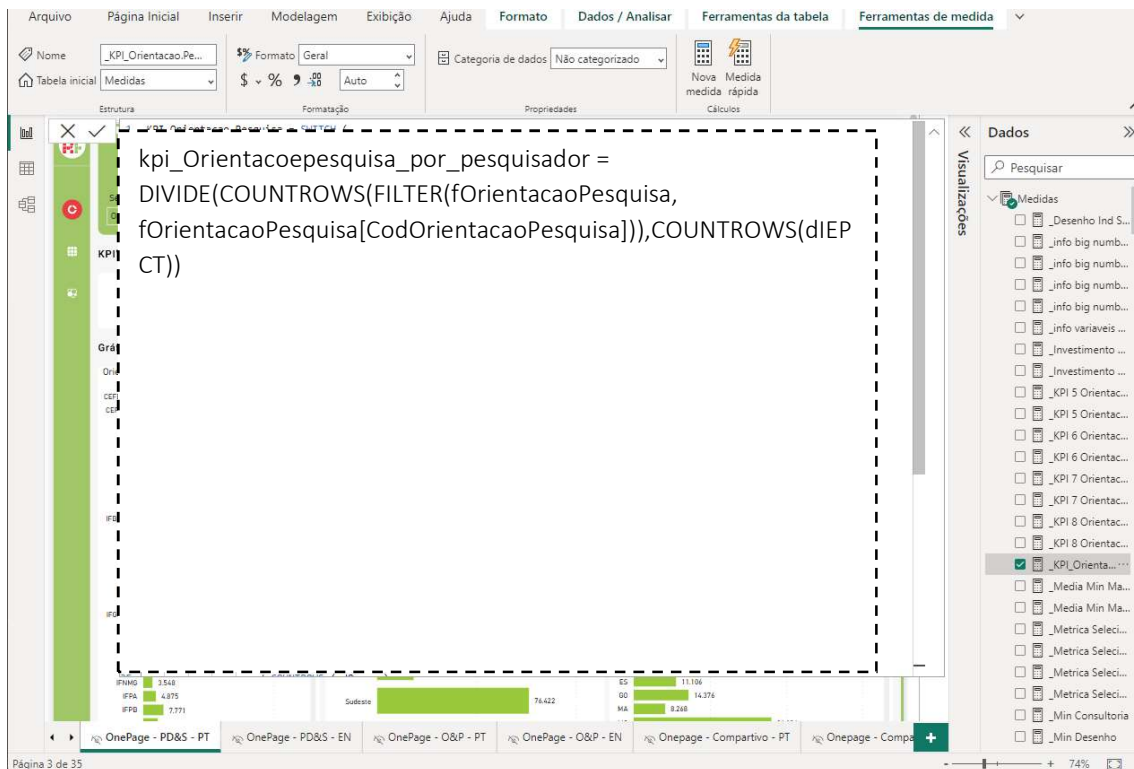


Figura B. 12: Ilustração da implementação da codificação de um KPIs no *Data Warehouse* do Power BI com DAX



Figura B. 13: Ilustra a página inicial que contém a seleção das duas Tecnologias Desenvolvidas pelo Grupo de Pesquisa NUMDI: esquerda registro BR 51 2023 001019 7 (INPI, o Autor, 2023) e a direita registro BR 51 2023 002411 2 (INPI, Sgarbi, 2023).



Figura B. 14: Ilustração do UX design do MVP proposto neste estudo registrado sob o registro BR 51 2023 001019 7 (INPI)

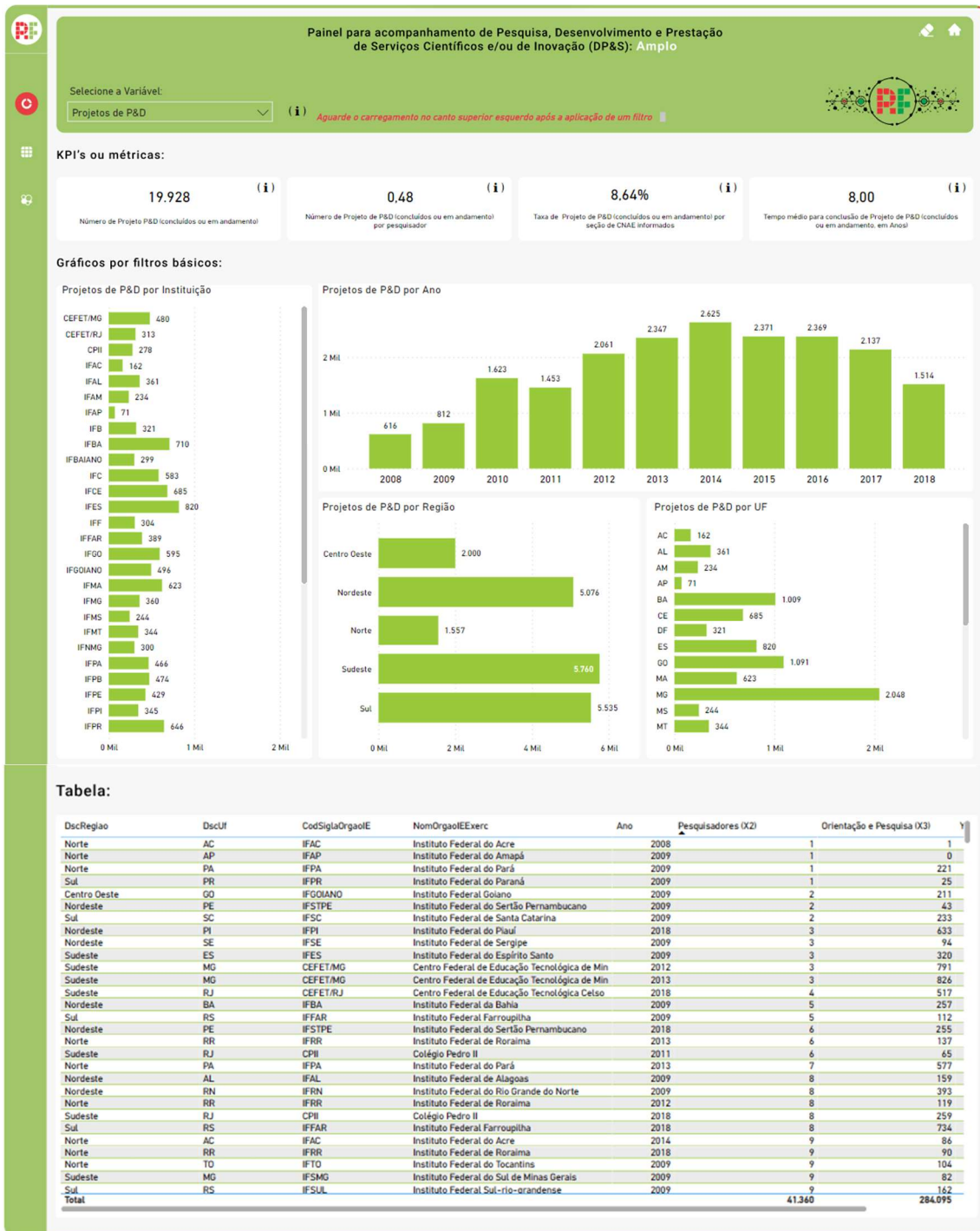


Figura B. 15: Ilustração do UX design dashboard para o módulo – Amplo



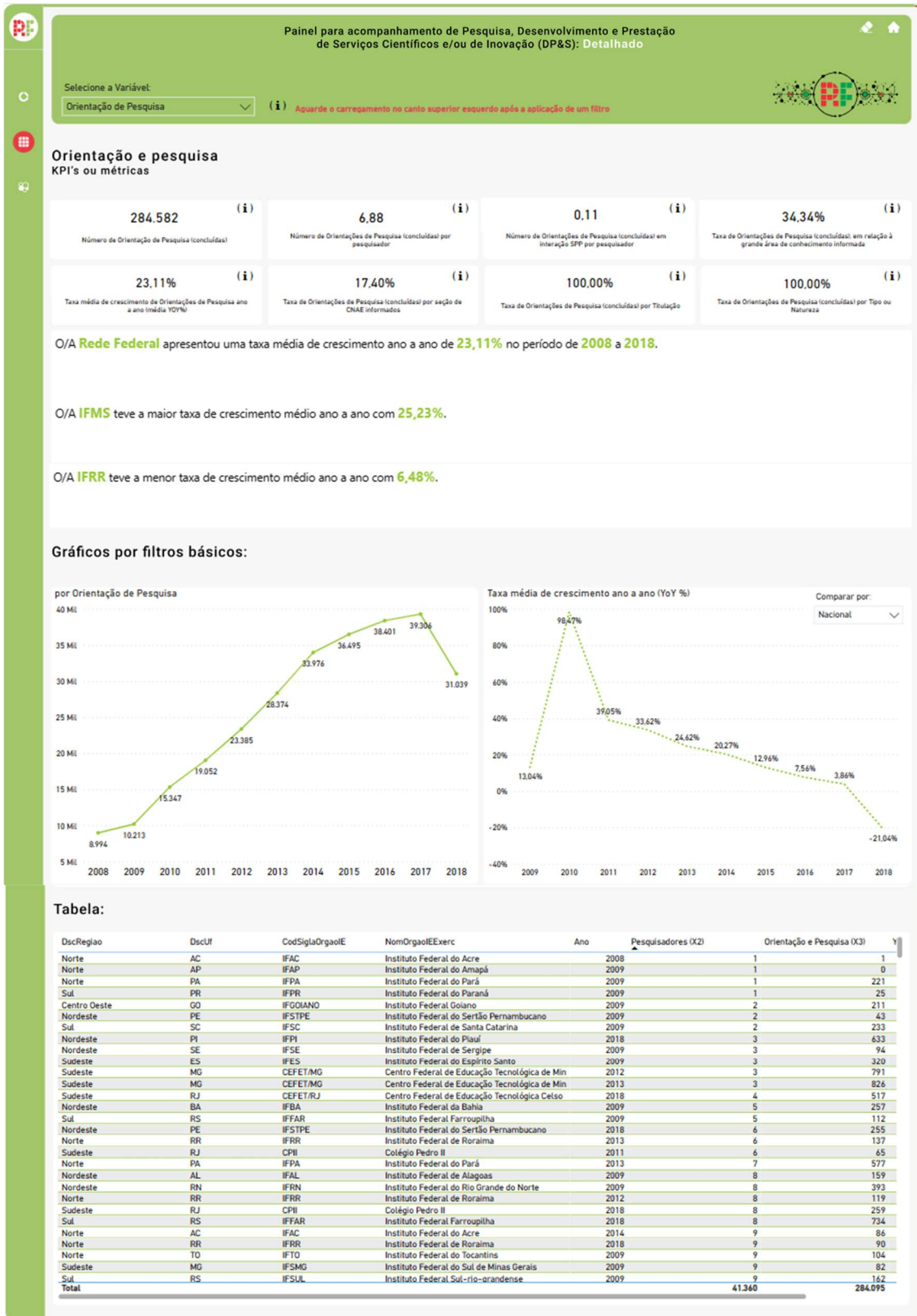


Figura B. 16: Ilustração do UX design dashboard para o módulo – **Detalhado**

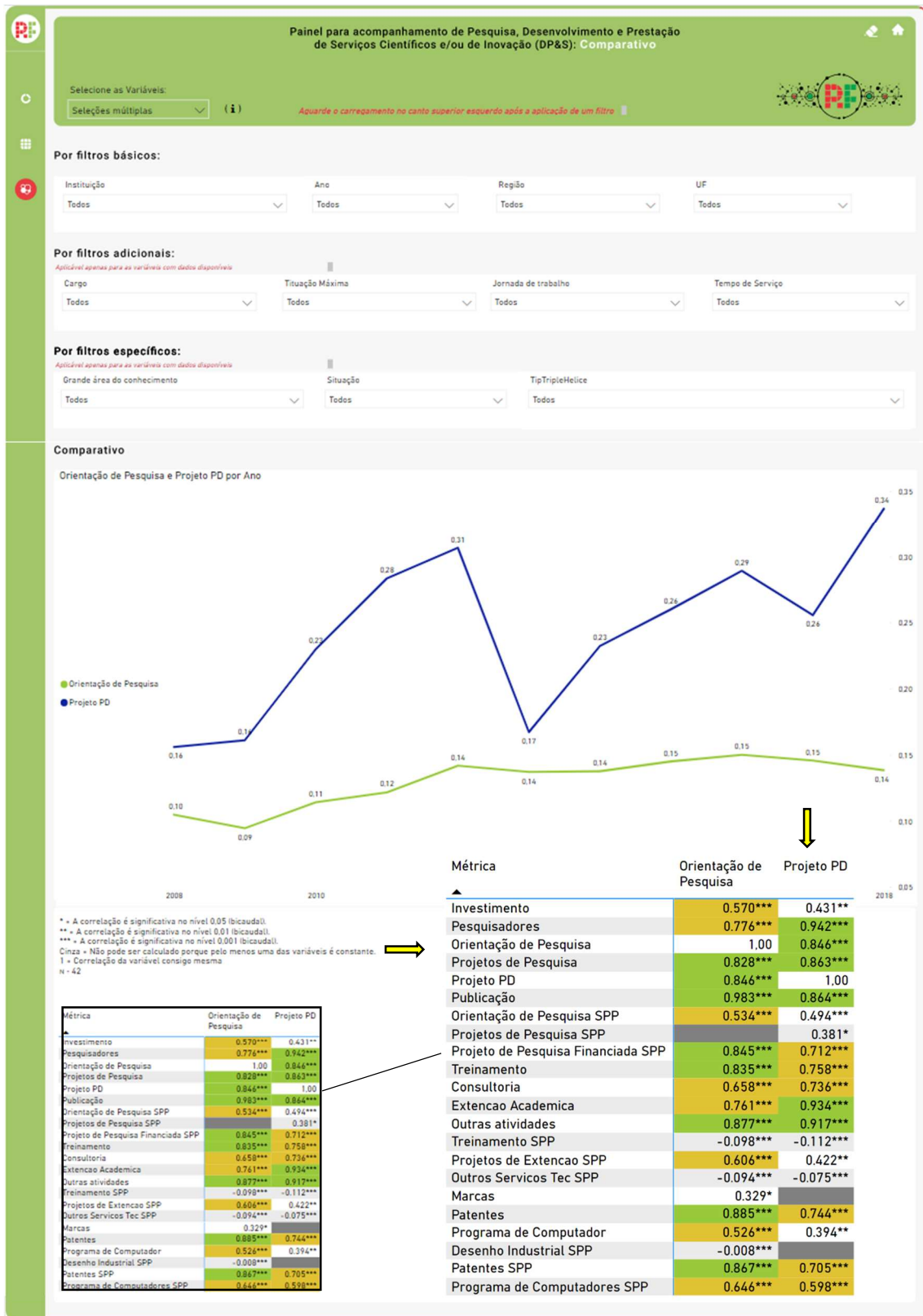
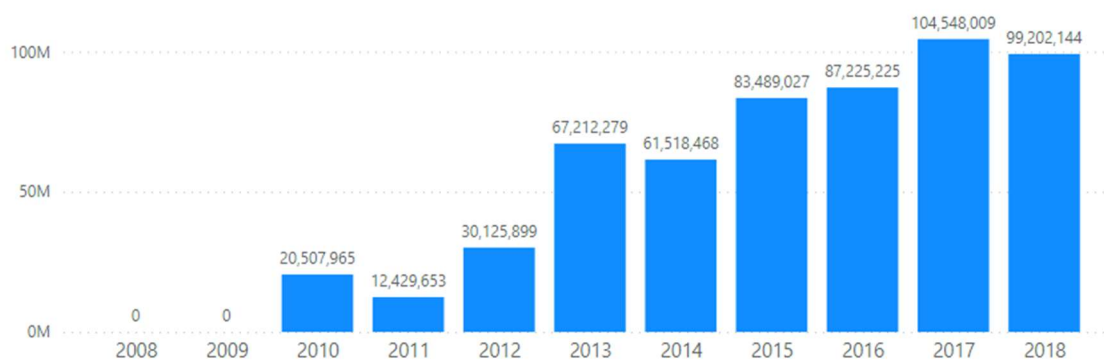


Figura B. 17: Ilustração do UX design dashboard para o módulo – Comparativo

Gráfico B. 1: Evolução dos investimentos em P&D entre 2008 e 2018 na RFEPCT



Fonte: dados disponíveis no micro dados do IPEA, Censo da educação Superior, organizado e disponível no MVP proposto nesta tese.

Nota: Filtrado por **Ano** (quando não é em branco), **Métrica - X<sub>1</sub>** (é Investimento federal em P&D), **Cálculo** (é Soma), **Ano** (não é em branco).

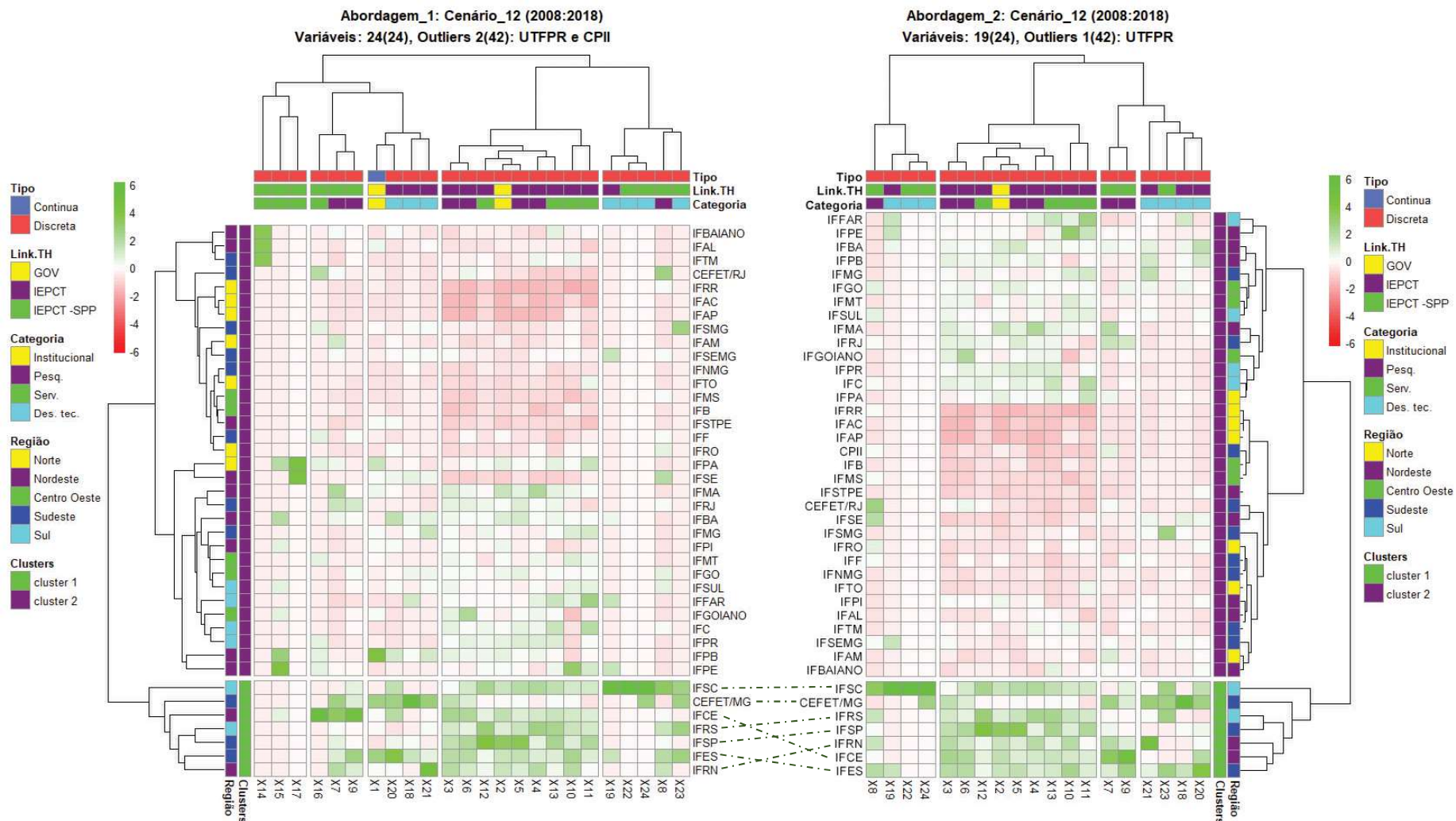


Figura B. 18: Estudo comparativo de Agrupamento Hierárquico das Abordagens 1e 2 para o intervalo temporal entre 2008 e 2018 da RFEPC.T.

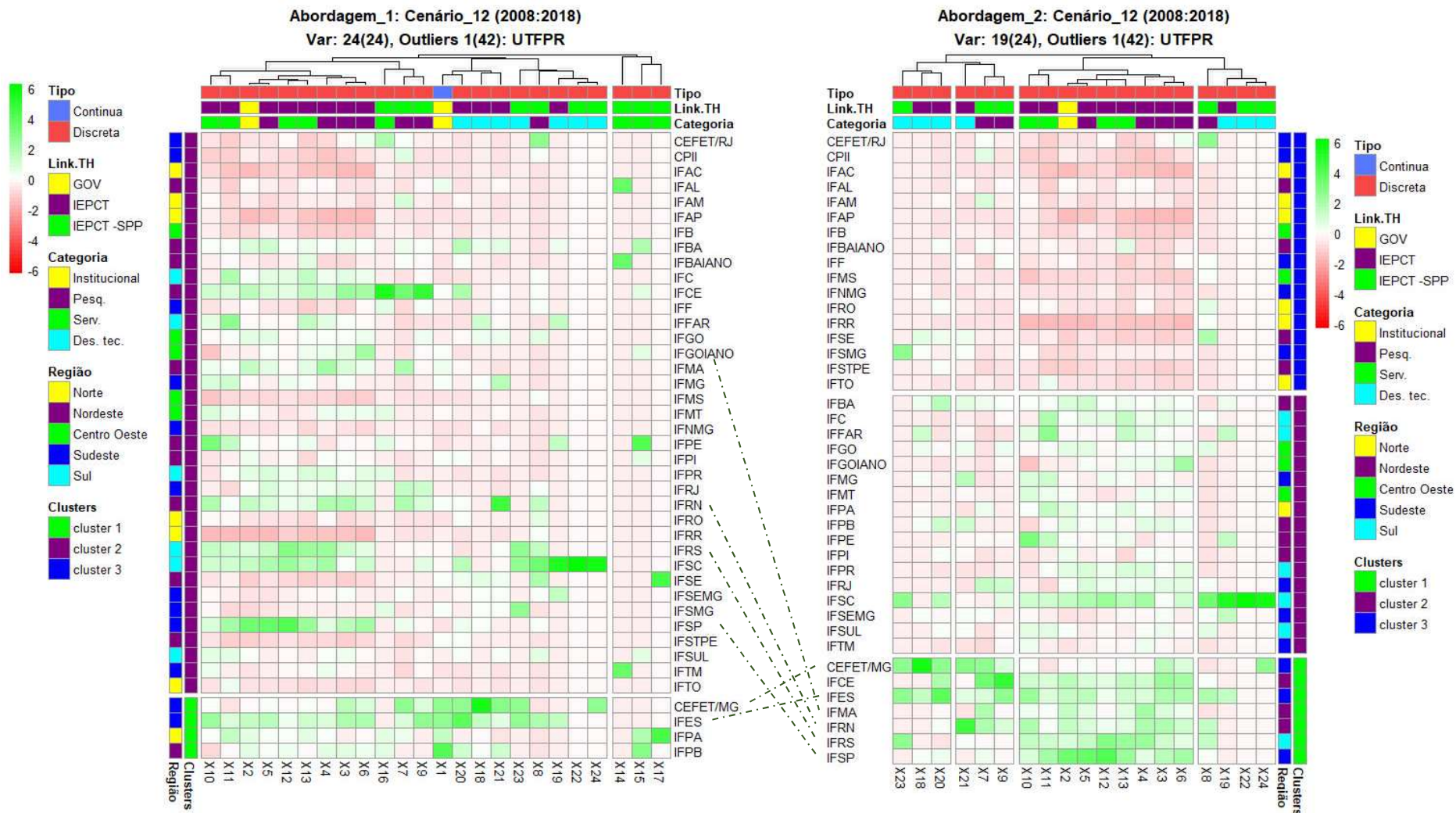


Figura B. 19: Estudo comparativo de Agrupamento Não Hierárquico (NHC: *K-means*) das Abordagens 1e 2 para o intervalo temporal entre 2008 e 2018 da RFEPC

## Apêndice C - Métodos de Interdependência Multivariada Exploratória ou Mineração de Dados

Os Métodos de Interdependência Multivariada Exploratória, subdividem-se em técnicas de Análise de Agrupamento ou *Clusters*, Análise Fatorial, Análise por Correspondência e Escalonamento Multidimensional, entre outros (Fávero *et al.*, 2020; Hair *et al.*, 2019). As análises a seguir foram realizadas utilizando a linguagem R.

O Método de Mineração de Dados (em inglês, *Data Mining* – DM) é um processo que busca extrair conhecimento de grandes conjuntos de dados (Han *et al.*, 2012), sendo os Métodos de Interdependência Multivariada Exploratória como um dos principais tipos utilizados na identificação de padrões, tendências e ou anomalias dos dados.

### C.1 - Análise *Cluster* Hierárquico (HC) e Não Hierárquico (NHC).

A análise de *clusters* busca identificar conjuntos de dados com atributos similares, subdividindo-se em Agrupamento Hierárquico (em inglês, *Hierarchical Cluster* – HC) e Agrupamento não Hierárquico (em inglês, *Non Hierarchical Cluster* – NHC) (Fávero *et al.*, 2020; Garson, 2014; Hair *et al.*, 2019; Zhang *et al.*, 2017)

Essa popular análise multivariada, exploratória, visa identificar grupos com características (dis)semelhantes. Esses dividem-se em hierárquicos e não hierárquicos.

Inicialmente foi elaborada uma HC, que é um método quantitativo para classificação de dados e pode ser vista como um tipo de taxonomia numérica. Essa taxionomia é representada sob a forma de Dendrogramas. Esses representam imagetivamente a distância hierárquica dos dados relacionados às observações (organizados em linhas), a partir de suas muitas dimensões ou variáveis (organizadas em colunas). Por fim, as linhas (observações ou *clusters*) são contrastadas com as colunas (dimensões ou variáveis) calculando-se a distâncias entre os valores dessas variáveis, através de funções matemáticas que medem suas distâncias ou (dis)similaridades.

Essas são usadas para determinar quão próximos ou distantes dois pontos (ou observações) estão entre si no espaço multivariado. Aqui estão algumas das distâncias ou métricas de dissimilaridade mais comuns utilizadas (Fávero *et al.*, 2020; Garson, 2014; Hair *et al.*, 2019; Zhang *et al.*, 2017).

A **Distância Euclidiana** trata-se, simplificada, da distância "em linha reta" entre dois pontos no espaço, e é a mais comumente usada. É calculada pela raiz quadrada

da soma dos quadrados das diferenças entre as coordenadas dos pontos, também conhecida como distância quadrática.

A **Distância de Manhattan (ou Distância *City-block*)** é a soma das diferenças absolutas entre as coordenadas dos dois pontos.

A **Distância de Minkowski** é a generalização das Distâncias Euclidiana e de Manhattan. A distância de Minkowski de ordem  $p$  entre dois pontos é a  $p$ -ésima raiz da soma das  $p$ -ésimas potências das diferenças entre suas coordenadas. A Distância Euclidiana e de Manhattan são casos especiais de  $p = 2$  e  $p = 1$ , respectivamente.

A **Distância de Mahalanobis** considera a correlação entre as variáveis e é a distância entre pontos no espaço multivariado quando essas estão correlacionadas.

A **Distância de Jaccard** é principalmente aplicada a dados binários, esta medida é o complemento do coeficiente de Jaccard, que é a proporção de presenças comuns aos dois objetos em relação ao número total de presenças.

A **Distância de Cosseno** que mede o cosseno do ângulo entre dois vetores, sendo amplamente utilizado em Análise de Texto (em inglês, *Text Mining*).

A **Distância de Correlação** que se baseia na correlação de Pearson entre dois pontos.

Como técnica computacional, a HC é calculada sob forma de uma matriz fenotípica de elementos dendrodistantes  $d_{ij}$  (Sneath *et al.*, 1963; Sokal, 1958; 1966), considerando uma matriz  $X$  de dados brutos com  $t$  linhas (ou etiquetas ou objetos) e  $n$  colunas. Primeiro, é importante considerar se os dados estão dentro de uma mesma unidade de medida; se não estiverem, faz-se necessário padronizá-los a fim de garantir consistência e comparabilidade nas análises subsequentes. Os dados  $n$  são padronizados por uma matriz  $Z$  a partir de uma matriz de dados original  $X$  com  $t$  linhas (observações ou etiquetas) e  $n$  colunas.

A padronização de dados é um processo crucial no seu pré-processamento para muitas técnicas estatísticas e de aprendizado de máquina. O objetivo é padronizar os dados de modo que eles tenham propriedades específicas e uma unidade adimensional entre eles capaz de comparar unidades distintas como, por exemplo, metros, litros, valores etc., muitas vezes tornando-os mais amigáveis para os algoritmos. Aqui estão os tipos mais comuns de padronização de dados:

O **Z-Score** (em inglês, *Standard Score*) cada valor de uma característica (ou variável) é subtraído da média ( $\mu$ ) dessa característica e dividido pelo desvio padrão ( $\sigma$ )

dessa característica. O resultado é que a característica padronizada terá uma  $\mu$  de 0 e um  $\sigma$  de 1.

A transforma os dados para que eles se ajustem a um intervalo específico, geralmente [0,1] ou [-1,1]. Para fazer isso, subtrai-se o valor mínimo da característica e divide-se pela diferença entre o valor máximo e o valor mínimo.

A **Escala Robusta** é semelhante à padronização por *Z-score*, mas utiliza a mediana e o intervalo interquartil (IQR) em vez da média e do desvio padrão. Isso torna a padronização mais resistente a *outliers*.

A **Normalização de Vetores** é principalmente adotada quando os dados estão organizados em formato de vetor (como em análise de texto), onde cada vetor (ou registro) é dividido por sua magnitude, resultando em um vetor unitário.

A **Binarização** padroniza os dados definindo um valor limiar. Todos os valores acima desse limiar são transformados em 1 e todos os valores abaixo ou iguais em 0.

O **One-hot Encoding** ou **Duming** aplicado principalmente para dados categóricos, em que cada valor único da característica é padronizado em uma nova característica binária.

A **Transformações Logarítmicas** e de **Potência** são usadas para tornar os dados mais normalmente distribuídos, o que pode ser útil para algoritmos que assumem que os dados têm uma distribuição normal.

A escolha do tipo de padronização depende da natureza dos dados e dos requisitos do algoritmo ou técnica estatística que será aplicada subsequentemente. Em muitos casos, é útil experimentar várias técnicas de padronização para determinar qual delas oferece o melhor desempenho para um determinado problema. Para este estudo foi adotada a padronização **Z-Score**.

Detectar *outliers* é crucial em muitas análises estatísticas, incluindo a HC, uma vez que sua presença pode afetar significativamente os resultados. Várias técnicas podem ser aplicadas para detectar *outliers*, dependendo de estarem analisando uma única variável (univariada) ou múltiplas variáveis ao mesmo tempo (multivariada), considerando os valores de  $n$ . As técnicas de Análise univariada mais comum são:

- O **Gráficos de Box-Plot** (*Box and Whisker Plot*) que é uma representação gráfica para descrever a distribuição de dados. Os pontos fora dos "bigodes" do gráfico são frequentemente considerados *outliers*.



- O **Gráficos de Histograma** permite visualizar a distribuição de frequências de um conjunto de dados e identificar possíveis *outliers* visualmente.
- O **Z-Score** permite visualizar os valores considerados levados (por exemplo, maior que 3 ou menor que -3), que podem ser considerados possíveis *outliers*.
- O **IQR (Intervalo Interquartil)** permite identificar visualmente qualquer valor que esteja abaixo do 1º quartil menos 1,5 vezes o IQR ou acima do 3º quartil mais 1,5 vezes o IQR pode ser considerado um *outlier*.

As técnicas de Análise multivariada mais comum são:

- O **Gráfico de Dispersão (Scatter Plot)** com Cores permite visualizar múltiplas variáveis e colorir pontos com base em outra variável ou métrica pode ajudar a identificar *outliers* multivariados.
- A **Análise Fatorial Exploratória por Componentes Principais (Exploratory Factor Analysis by Principal Components - PCA)** pode ser usada para reduzir a dimensionalidade dos dados, e os componentes resultantes podem ser usados para identificar *outliers*.
- A **Distância de Mahalanobis** é uma medida de distância que leva em consideração a correlação entre as variáveis. Valores com uma Distância de Mahalanobis muito grande podem ser considerados *outliers* multivariados. O algoritmo comumente adotado para essa análise é o *Blocked Adaptive Computationally-efficient Outlier Nominators (BACON)* (Billor *et al.*, 2000; Fávero *et al.*, 2020, p.375).
- O **Métodos de Clustering**, como o DBSCAN pode identificar grupos densos de pontos e classificar pontos que estão longe desses grupos como *outliers*.
- O **Métodos de Estimação Robusta** que busca modelar os dados enquanto é resistente aos efeitos de *outliers*.

Ao identificar possíveis *outliers* usando as possíveis técnicas acima, é importante lembrar que a sua presença não é necessariamente um erro. Elas podem representar

variações genuínas nos dados ou podem ser indicativos de erros de medição ou de registro. Portanto, a decisão de tratar ou remover outliers deve ser feita com cuidado e considerando o contexto do estudo (Fávero *et al.*, 2020; Field, 2018; Hair *et al.*, 2019).

Na prática, os dados brutos da Tabela B. 3 (Apêndice B) podem ser representados por uma matriz  $X$  composta de  $n = 24$  variáveis, por exemplo, Investimento público em Pesquisa e Desenvolvimento ( $X_1$ ), Pesquisadores ( $X_2$ ) e Orientações de Pesquisa ( $X_3$ ) e  $t = 42$  ICTs (ou etiquetas ou observações). Portanto, como cada dado deve ser equivalente a outros dados brutos da matriz  $X$ , com  $n$  variáveis (ou colunas) e rótulos  $p$  (ou linhas), cada dado original  $X_i$  deve ser transformado em  $Z_i$  padronizado após os cálculos (Everitt *et al.*, 2011; Hair *et al.*, 2019). Assim, todo  $Z_i$  pode ser descrito por:

$$Z_i = \frac{X_i - \bar{X}_i}{\sigma_x}$$

Onde  $\bar{X}_i$  é a  $\mu$  e seu  $\sigma_x$  correspondente, pode ser definido como (Everitt *et al.*, 2011):

$$\sigma_x = \sqrt{\frac{\sum_{i=1}^p (X_i - \bar{X}_i)^2}{p}}$$

Considerando  $p$  observações dos dados originais de  $X_i$ .

Padronizados os dados para  $Z_i$ , aplicou-se a métrica euclidiana, que associa dois elementos padronizados,  $Z_i$  e  $Z_j$ , de um conjunto de dados, caracterizado por  $t$  linhas e  $n$  colunas (Fávero *et al.*, 2020; Field, 2018; Hair *et al.*, 2019). Essa padronização tem o efeito de minimizar as diferenças do *cluster* a ser formado, pois eles serão bem separados em relação a variável  $Z_i$  porque apresentam uma variância grande o suficiente  $\sigma_x^2$  para os diferenciarem (Everitt *et al.*, 2011).

Este procedimento simplifica e reduz o conjunto de dados multidimensionais, promovendo uma medida razoável das distâncias euclidianas  $d_{ij}$  de diferentes elementos (Sneath *et al.*, 1963; Sokal, 1958; 1966). Então, um algoritmo simples começa a encontrar dados que estão mais próximos uns dos outros com base na distância euclidiana. Os resultados desse procedimento serão utilizados para a construção do dendrograma. Assim, todo  $d_{ij}$  pode ser descrito por:

$$d_{i,j} = \sqrt{\sum_{k=1}^p (Z_{i,k} - Z_{j,k})^2}$$

Calculada as distancias euclidianas  $d_{i,j}$  o próximo passo é calcular as distâncias entre os *clusters* a serem formados por esses pontos, ou seja, sua subsequente aglomeração ou divisão, dependendo se está aplicando uma abordagem aglomerativa ou divisiva.

A Abordagem Aglomerativa inicia o processo de separação dos *clusters* considerando cada ponto como um *cluster* individual. Então, os dois *clusters* mais próximos são combinados em um único, e esse processo é repetido até que todos os pontos estejam em um único *cluster*. Para determinar quais *clusters* combinar em cada etapa, utilizou-se um critério de ligação (ou método de ligação) para definir a distância entre dois *clusters*. Os métodos de ligação mais comuns (Fávero *et al.*, 2020; Field, 2018; Hair *et al.*, 2019):

- A **Ligação Simples** (Mínima) considera que a distância entre dois *clusters* é a distância mínima entre qualquer par de pontos, um de cada *clusters*.
- A **Ligação Completa** (Máxima) considera que a distância entre dois *clusters* é o inverso da Ligação Simples.
- A **Ligação Média** considera que distância entre dois *clusters* é a média das distâncias entre todos os pares de pontos, um de cada *clusters*.
- A **Ligação do Centroide** considera que a distância entre dois *clusters* é a distância entre os seus centroides.
- A **Ligação de Ward** (ou Ward.D2) combina os *clusters* que resulta no mínimo aumento na soma total dos quadrados dentro de cada *clusters*.

A Abordagem Divisiva inicia com todos os pontos em um único *cluster* e, em seguida, divide-se o *cluster* em outros menores a cada etapa, até que cada ponto esteja em seu próprio *cluster*.

Neste estudo será adotada a Abordagem Aglomerativa utilizando o Método Ward.D2. Deve-se notar que devido ao uso do Método de Ligação de Ward, a magnitude não representa a distância euclidiana  $d_{i,j}$  entre os pares, a exemplo quando se utiliza os métodos de agrupamento simples, completo ou médio. No método de ligação Ward.D2, adotado, a distância representa o aumento ponderado na soma dos quadrados intra-*cluster*

resultante da fusão de dois *clusters*. Esta distância é calculada pela seguinte equação (Murtagh *et al.*, 2014, p. 292):

$$D^2(A, B) = \frac{|A| * |B|}{|A| + |B|} * d(A, B)^2$$

onde  $D^2(A, B)$  representa a distância ao quadrado entre dois *clusters*  $A$  e  $B$ . A expressão  $\frac{|A|*|B|}{|A|+|B|}$  é o fator de ponderação, onde  $|A|$  e  $|B|$ , indicam o número de observações nos *clusters*  $A$  e  $B$ , respectivamente. Por fim,  $d(A,B)^2$  denota a distância ao quadrado entre os centros (ou médias) dos *clusters*  $A$  e  $B$ . A distância entre estes centros pode ser determinada através de várias métricas, com a distância euclidiana sendo uma das mais comumente utilizadas. O propósito dessa fórmula é minimizar a soma dos quadrados *intra-cluster* e, ao mesmo tempo, maximizar a diferença entre eles.

Uma vez que os *clusters* são formados, é comum visualizar a estrutura de agrupamento usando um dendrograma, que é um diagrama em árvore mostrando o processo de aglomeração ou divisão e as distâncias (ou alturas) em que diferentes *clusters* são combinados ou divididos.

O dendrograma é um diagrama que exhibe aglomerados formados por observações de agrupamento em cada etapa e em seus níveis de similaridade. As distâncias de agrupamento foram calculadas ao longo do eixo vertical e as diferentes observações são listadas ao longo do eixo horizontal. Este estudo é utilizado para definição da quantidade de *clusters* a ser considerado no método não hierárquico e para identificar a taxionomia ou estrutura das observações, a partir dos variáveis consideradas, conforme destacam Fávero *et al.* (2020).

Concluída a análise do dendrograma, Fávero *et al.* (2020, p. 344) enfatizam a importância de verificar se os *clusters* identificados apresentam máximas diferenças entre si. Segundo ele, esta etapa é tão crucial quanto a escolha do método de distância e agrupamento. Neste sentido, ele sugere a aplicação da ANOVA *one way*, não para testar hipóteses, mas para interpretar o resultado descritivo dos *clusters* identificados. O teste F auxilia, portanto, na validação e determinação de quais variáveis não são estatisticamente significativas para diferenciar os *clusters* formados. Variáveis com  $p\text{-valor} > 0,05$  são consideradas não relevantes para a separação deles. Se essas variáveis forem

identificadas, recomenda-se refazer a análise HC sem elas e comparar os resultados antes e depois da exclusão.

Fávero *et al.* (2020, p. 338-344), esclarece que a decisão sobre quais variáveis contribuem para a separação dos *clusters* deve ser baseada no contexto da análise e no objetivo do estudo. Assim, cabe ao pesquisador determinar sua inclusão ou supressão, considerando o conhecimento do ambiente de negócio analisado.

Nesta primeira etapa que envolveu a análise dos dados a partir do algoritmo HC, as observações são agrupadas sequencialmente com o objetivo de formar *clusters* internamente homogêneos e, ao mesmo tempo, distintos (heterogêneos) entre si.

Em seguida, o algoritmo NHC foi aplicado, cujo objetivo é identificar *cluster* onde as distâncias entre as observações pertencentes a ele sejam minimizadas (ou seja, o mais homogênea possível) e as distâncias entre os diferentes *clusters* sejam maximizadas (ou seja, o mais heterogênea possível), solução similar à adotada pela HC utilizando o método de agrupamento Ward.D2. Embora compartilhe algumas semelhanças com o método HC, ele adota uma abordagem distinta. Em vez de se concentrar nas distâncias entre as observações em si, ele foca em como elas são homogêneas dentro de um *cluster* específico. Os principais métodos de agrupamento NHC incluem (Abd-Almageed *et al.*, 2006; Ester *et al.*, 1996; Fávero *et al.*, 2020; Park *et al.*, 2009):

- O ***K-means*** busca minimizar a variância dentro do *cluster* e maximizar a variância entre os *clusters*. O número de *clusters* (*k*) deve ser especificado a priori, advindo de um método como Método do Cotovelo (em inglês, *Elbow Method*) (Thorndike, 1953), da Silueta (em inglês, *Silhouette*) (Rousseeuw, 1987), Método da Soma dos Quadrados (Dendrograma), entre outros.
- O ***K-medoids*** (em inglês, *Partitioning Around Medoids - PAM*): Semelhante ao *K-means*, mas, em vez de calcular a média dos pontos em um cluster, usa um dos pontos reais do *dataset* como referência ou "medióide".
- O **DBSCAN** (*Density-Based Spatial Clustering of Applications with Noise*): Ao contrário dos métodos baseados em partições, como o *K-means*, o DBSCAN é um método baseado em densidade que pode identificar qualquer número de clusters de qualquer forma, baseando-se na densidade de pontos.

- O Método de Mistura Gaussiana (em inglês, *Gaussian Mixture Model - GMM*): Este é um método probabilístico que assume que os dados são gerados a partir de várias distribuições gaussianas diferentes.

Estes são apenas alguns dos métodos mais comuns. Existem muitos outros algoritmos e variações utilizando o método de agrupamento NHC. A escolha do método mais apropriado está associada a natureza dos dados, ao objetivo da análise e conhecimento e domínio do modelo de negócio analisado.

Considerando a natureza dos dados, ao objetivo da análise e conhecimento e domínio do modelo de negócio analisado, neste trabalho optou-se por utilizar o método *K-means*. O método de agrupamento NHC busca reduzir a possível subjetividade humana em identificar possíveis agrupamentos naturais. Esse método de agrupamento, não supervisionado, divide os dados em um número específico de *k-clusters* para comprimir ou agrupar casos originais. Neste trabalho, o número de *k-clusters* (*k*) utiliza *a priori*, o resultado alcançado pelo Método da Soma dos Quadrados (Dendrograma).

Seu algoritmo matemático preciso determina que cada grupo seja nucleado por um *centroide*, o ponto central do *cluster* (Everitt *et al.*, 2011; Hair *et al.*, 2019). De acordo com Macqueen, James (1967), *K-means* representa uma generalização da média amostral comum, onde o centroide é um *datum* escolhido aleatoriamente. O critério de particionamento é uma medida de distância, conforme explicado abaixo (Macqueen, James, 1967).

O critério de particionamento refere-se ao método usado para designar cada ponto de dados a um *cluster* específico. No algoritmo *K-means*, essa decisão é tomada com base na distância entre um ponto de dados e o centroide de cada *cluster*. Geralmente, com dados padronizados, utiliza-se a distância euclidiana para calcular essa distância (conforme detalhado anteriormente), mas outras métricas também são opções viáveis. O objetivo é minimizar a soma das distâncias quadradas entre os pontos de dados e o centroide de seu cluster atribuído.

Após o cálculo das distâncias euclidianas  $d_{i,j}$ , o próximo passo é determinar as distâncias entre os *clusters* formados a partir desses pontos. Para isso, é aplicado um algoritmo específico, um procedimento que se assemelha ao utilizado em métodos HC. Conforme visto acima, existem diversos métodos de agrupamento NHC disponíveis. Ao método adotado *K-means* alguns algoritmos são propostos para sua otimização. O

Quadro C. 1 a seguir apresenta os algoritmos mais comuns.

Quadro C. 1: Comparação entre Algoritmos de Otimização para *K-means*

<b>Critério</b>	<b>Lloyd-Forgy</b>	<b>MacQueen</b>	<b>Hartigan e Wong</b>
<b>Referencial</b>	(Lloyd, 1982)	(Macqueen, James, 1967)	(Hartigan, 1975)
<b>Inicialização</b>	Seleciona $k$ pontos aleatoriamente do conjunto de dados como centroide.	Seleciona $k$ pontos aleatoriamente do conjunto de dados como centroide.	Métodos mais sofisticados. Pode usar inicialização aleatória, mas também explora reatribuições iniciais para otimizar a soma dos quadrados <i>intra-clusters</i> .
<b>Atribuição</b>	Em cada iteração, cada ponto é atribuído ao cluster cujo centroide é mais próximo usando a distância Euclidiana.	A cada ponto lido, é atribuído ao cluster cujo centroide é mais próximo. Não espera por uma iteração completa para reatribuir.	Atribuição é baseada na minimização da soma total dos quadrados <i>intra-cluster</i> . Pode reatribuir pontos várias vezes durante uma única iteração para otimizar a soma.
<b>Atualização</b>	O centroide de cada cluster é recalculado tomando a média dos pontos atribuídos a ele após cada iteração completa.	O centroide é atualizado incrementalmente a cada ponto atribuído. Isso pode ser feito usando uma média ponderada.	Recalcula o centroide de forma semelhante ao Lloyd-Forgy, mas com considerações adicionais para minimizar a soma dos quadrados.
<b>Convergência</b>	Converge quando os centroides não mudam entre duas iterações consecutivas ou a mudança é menor que um limiar pré-definido.	Pode ter problemas de convergência em algumas situações devido à sua natureza incremental. Pode requerer mais iterações para convergir.	Utiliza uma técnica de troca de pontos entre clusters para garantir uma convergência mais eficiente. Converge quando a redução na soma total dos quadrados se torna insignificante ou nula.

Neste trabalho, em termos matemáticos, para um *cluster*, utilizando o método de agrupamento *K-means* e aplicando o algoritmo de otimização Hartigan e Wong, contendo pontos  $x_1, x_2, \dots, x_n$ , o centroide  $C$  foi calculada como:

$$C = \frac{1}{n} \sum_{i=1}^n x_i$$

onde  $n$  é o número total de observações ou elementos no conjunto de dados,  $x_i$  representa a  $i$ -ésima observação ou elemento do conjunto de dados e  $C$  é a posição do centroide do *cluster*.

O centroide é ajustado a cada iteração do algoritmo, com cada *datum* sendo associado ao *cluster* cujo centro está mais próximo. Este centro se adapta à medida que novos dados são agregados ao *cluster*. Uma vez que todos os pontos tenham sido atribuídos a um *cluster*, o centroide torna-se o ponto médio desses pontos. O processo de designação de pontos com base na distância ao centroide e sua subsequente recalibração continua até que elas se estabilizem, ou até que um limite predefinido de iterações seja atingido, procedimento esse adotado neste estudo utilizando. A Figura C. 1 apresenta o procedimento transcorrido para alcançar o resultado inicial NHC, adotando como referência os dados utilizados no Cenário 12, na Abordagem 2 (Subseção 3.5.9.1).





## C.2 – Análise Fatorial Exploratória (EFA) por Componentes Principais (PCA)

A Análise Fatorial Exploratória (em inglês, *Exploratory Factor Analysis* - EFA) por Componentes Principais (em inglês, *Principal Components* - PCA) é uma técnica de álgebra linear multivariada. Ela é recomendada para reduzir a dimensionalidade, entender a estrutura dos dados, elaborar indicadores e verificar se as variáveis latentes ou subjacentes ou *clusters* explicam satisfatoriamente as variáveis originais a partir de seus coeficientes de correlação (Fávero *et al.*, 2020; Field, 2017; Hair *et al.*, 2019; Hotelling, 1933; Kim *et al.*, 1978; Pearson, 1896; Spearman, 1904).

Essas variáveis latentes (VL) são denominadas de Fatores ou Componentes Principais, dependendo da abordagem, se realizada um EFA ou PCA, respectivamente.

Aplicando-se a abordagem EFA as variáveis latentes são denominadas de "Fatores". Esta abordagem procura identificar fatores subjacentes que explicam as correlações ou padrões entre um conjunto de variáveis observadas. Por exemplo, se aplicasse um questionário sobre interação ICTs-SPP com perguntas relacionadas à eficácia, comunicação e satisfação, uma EFA poderia revelar que existe um "Fator" subjacente relacionado à qualidade da interação que influencia todas essas três áreas.

Aplicando-se a abordagem PCA as variáveis latentes são denominadas de "Componentes Principais". Ao contrário da EFA, a PCA é uma técnica de redução de dimensionalidade que transforma as variáveis originais em um novo conjunto de variáveis (os componentes principais) que são combinações lineares das originais. A principal diferença é que a PCA busca maximizar a variação total dos dados, enquanto a EFA busca explicar a estrutura de correlação entre variáveis. Por exemplo, se possuísse um conjunto de dados sobre os resultados dos canais de PD&S em diferentes ICTs, a PCA pode identificar um ou mais componentes principais que representem uma tendência geral de desempenho dos canais de PD&S comuns a todas as ICTs.

Essas VL tendem a ser numericamente inferiores a quantidade e a retenção máxima de informações das variáveis originais. Cada VL representa uma combinação linear das variáveis originais padronizadas e não são correlacionados entre si, para mais detalhes consultar Kim *et al.* (1978).

A primeira VL, normalmente, tende a representar o maior percentual de explicação da variância total das variáveis originais; a segunda, o segundo maior percentual e assim sucessivamente. As VL representam um conjunto de variáveis específicas que permitem sua interpretação e descrição.

Dentre os vários métodos utilizados na extração das VL, sua escolha essencialmente dependerá do objetivo da abordagem. Tinsley *et al.* (1987), Hair *et al.* (2019) e Fávero *et al.* (2020) explicam detalhadamente vários desses métodos em seus trabalhos. Fávero *et al.* (2020), destaca ainda que embora a abordagem EFA seja recorrente na literatura, em essência ela é uma técnica multivariada exploratória, ou de interdependência, sem caráter preditivo. Essa característica, demanda que caso novos dados ou variáveis sejam adicionados à base de dados original é necessário refazer a análise.

O primeiro passo da execução da EFA por PCA a partir das variáveis originais do estudo foi examinar a matriz de correlação Tabela B. 5: Análise das correções dos dados referentes as 24 variáveis relacionada aos canais de PD&S das 41 unidades da Rede Federal (Apêndice B). O segundo passo foi analisar se os dados Tabela B. 3(Apêndice B) eram adequados para proceder a EFA por PCA, para tal foram utilizados a Medida Kaiser-Meyer-Olkin de adequação de amostragem (KMO) e o Teste de Esfericidade de Bartlett para avaliar se os dados originais são adequados para prosseguir a análise.

A Medida de KMO (Kaiser, 1970) fornece a proporção da variância comum em relação a todas as variáveis analisadas. O Teste de Bartlett (Bartlett, 1954) compara a matriz de correlação com sua matriz identidade buscando rejeitar a hipótese nula de que a matriz de correlação seja estatisticamente igual à matriz identidade de mesma dimensão.

O Teste de Bartlett é considerado estatisticamente mais robusto, tornando seu resultado preferível ao da Medida de KMO (Fávero *et al.*, 2020). Por exemplo, mesmo que a Medida de KMO seja inferior a 0,5, se o Teste de Bartlett apresentar um  $p\text{-valor} \leq 0,05$ , considera-se que os dados originais são adequados para a análise. No entanto, se a Medida de KMO for  $\geq 0,5$ , mas o Teste de Bartlett indicar um  $p\text{-valor} \geq 0,05$ , a interpretação correta é que os dados originais não são adequados para seguir com a EFA por PCA.

O terceiro passo foi extrair as variáveis latentes, doravante denominadas de Componentes Principais. Para uma melhor interpretação desses resultados aplicou-se uma rotação (Fávero *et al.*, 2020; Field, 2017; Hair *et al.*, 2019; Kim *et al.*, 1978) à sua matriz. O objetivo da rotação é melhorar a interpretação dos resultados. Uma vantagem desse procedimento é que soluções padronizadas permitem estimar a contribuição única de cada Componente Principal.

Existem dois tipos de rotações: ortogonais e oblíquas. A rotação ortogonal é recomendada quando se busca identificar Componentes Principais que não estejam

correlacionados entre si, sendo útil, por exemplo, na criação de índices. Por outro lado, a rotação oblíqua é empregada quando é desejável que exista alguma correlação entre os componentes, não sendo aconselhável sua adoção caso um dos objetivos seja a criação de índices. Para verificar a robustez dos métodos, explorou-se diferentes rotações ortogonais na construção dos Componentes Principais. A normalização Kaiser (Varimax) proporcionou os melhores resultados, sendo neste trabalho, o método selecionado para rotacionar a matriz. Para mais detalhes sobre outros métodos, consultar a literatura específica (Field, 2017; Kim *et al.*, 1978).

O quarto passo foi determinar os autovalores e autovetores da matriz de correlação ou covariância. Estes representam a quantidade de variação capturada por cada Componente Principal e a direção dessa variação no espaço original, respectivamente. O quinto passo envolveu a análise dos autovalores para determinar quantos Componentes Principais apresentam uma variância representativa do conjunto de dados originais. Uma regra comum é selecionar Componentes Principais cujos autovalores sejam  $\geq 1$  (Fávero *et al.*, 2020; Field, 2017; Hair *et al.*, 2019). O sexto passo foi destacar os carregamentos mais fortemente associados a um determinado componente. Field (2017) sugere identificar aqueles que possuam magnitude  $\geq 0,4$ .

O sétimo passo envolve a interpretação dos Componentes Principais identificados, considerando os pesos (também chamados de cargas ou magnitudes) dos autovetores em cada Componente Principal para determinar quais variáveis têm maior influência sobre cada um. Se necessário, os Componentes Principais podem receber rótulos para facilitar a compreensão do que representam. Segundo Fávero *et al.* (2020) não há obrigatoriedade de nomear ou rotular. Para eles, o importante é que esses representem as novas variáveis explicativas (dependentes) apresentando poder explicativo, similar, as originais, tendo como benefício a redução da dimensionalidade e da não multicolinearidade, entre si. Esse resultado, como já mencionamos, permite sua aplicação em outras técnicas multivariadas preditivas que tenham essa premissa como restrição entre as variáveis explicativas, como por exemplo uma Análise de Regressão.

O oitavo passo é optativo, ele é utilizado quando no objetivo prever a criação de um índice. Para realizar esse passo os Componentes Principais derivados em análises subsequentes, como regressão, clusterização etc., são calculados. Essas pontuações (em inglês, *scores*) podem ser utilizadas para a construção de índices, por exemplo. Há vários métodos para extrair a pontuação dos Componentes Principais, os quais destacam-se o

método *Bartlett, Anderson-Rubin* e regressão (Fávero *et al.*, 2020; Field, 2017; Hair *et al.*, 2019). Neste Estudo foi adotado o método de regressão.

O método da regressão estima os coeficientes de pontuação ponderados ou (em inglês, *scores*) dos Componentes Principais. Uma outra vantagem é que essas pontuações podem ser correlacionadas mesmo quando os dos Componentes Principais são ortogonais. Para mais detalhes sobre outros métodos consultar (Fávero *et al.*, 2020; Field, 2017; Hair *et al.*, 2019).

Após a realização da PCA, é possível construir um índice denominado, para efeito didático de “Índice Baseado em Componentes Principais” (IBC), que é obtido a partir da soma ponderada dos Componentes Principais rotacionados. Esta ponderação é realizada considerando os percentuais das cargas dos Componentes Principais compartilhadas pelas variáveis originais. Deve-se notar que elas são adimensionais, uma vez que são derivadas considerando o *Z-Score* de cada variável original. Devido a sua natureza específica não é recomendável usar o IBC para fazer comparações percentuais entre diferentes casos ou unidades. No entanto, é viável determinar se um caso é superior ou inferior a outro com base em sua classificação relativa, conforme descrito por autores reconhecidos no campo.

Devido à natureza construtiva das cargas fatoriais, o valor resultante de uma análise fatorial não é recomendado ser utilizado para realizar comparações do tipo: o caso “A” apresentou um IBC percentualmente maior ou menor que o “B”. Porém, é possível estabelecer que o caso “A” apresentou um IBC maior ou menor que “B”, em relação ao seu posicionamento no ordenamento, conforme sugerem Fávero *et al.* (2017). Um processo genérico para criar um IBC é ilustrado a seguir.

Suponha-se que um pesquisador esteja interessado em analisar a “evolução<sup>34</sup>” de 5 “observações” (A, B, C, D e E) com base em 4 variáveis contínuas e/ou discretas ( $V_1$ ,  $V_2$ ,  $V_3$ ,  $V_4$ ). Caso haja variáveis fatoriais (exemplo, nominais) e/ou ordinais (exemplo, Escala *Likert*) é recomendado, por exemplo, realizar anteriormente uma Análise Fatorial e/ou por Correspondência, respectivamente. Estas variáveis foram coletadas para os anos ( $A_1$  e  $A_2$ ).

---

<sup>34</sup> O termo substantivo escolhido deve se alinhar ao contexto semântico e sintático da análise, bem como ao propósito da investigação. No cenário apresentado, se a intenção é expressar uma ideia similar a "performance", palavras como "desempenho", "eficácia", "produtividade", "resultados", "rendimento", "operacionalidade", "execução", "funcionamento", e "capacidade" podem ser empregadas. Por outro lado, se o foco recai sobre uma noção análoga a "evolução", termos como "progresso", "desenvolvimento", "crescimento", "melhoria", "ampliação", "aprimoramento", "otimização", "maturação", "refinamento" e "expansão" são alternativas adequadas.

A Tabela C. 1 e Tabela C. 2 apresentam os dados armazenados e tratados, após o desenvolvimento do processo de ETL (Subseção 3.5.3), em uma variável denominada “mtz\_numerica\_A1 e mtz\_numerica\_A2”, simulados para os anos A<sub>1</sub> e A<sub>2</sub>, respectivamente. A coluna  $ECP_i$  representa o escore do Componente Principal rotacionado;  $VC_i$  representa a Variância Compartilhada para cada Componente Principal rotacionado;  $IBC_n$  representa o Índice Baseado em Componentes Principais calculado, onde o  $n$  representa os dados de mtz\_numerica\_A<sub>n</sub>,  $IBC_n^{std}$  representa o  $IBC_n$  padronizado para uma escala entre [0,1] e “Order” representa o ordenamento ou posicionamento da observação em função do  $IBC_n^{std}$ , conforme sugere a literatura para uma melhor interpretação e clareza do índice.

Tabela C. 1: Dados Simulados para A<sub>1</sub> (em unidades arbitrárias)

<b>Id</b>	<b>V<sub>1</sub></b>	<b>V<sub>2</sub></b>	<b>V<sub>3</sub></b>	<b>V<sub>4</sub></b>	<b>ECP<sub>1</sub></b>	<b>VC<sub>1</sub></b>	<b>ECP<sub>2</sub></b>	<b>VC<sub>2</sub></b>	<b>IBC</b>	<b>IBC<sub>std</sub></b>	<b>Order</b>
<b>A</b>	150	10	15	30	,24	0,58	-,82	0,39	-0,18	0,51	4
<b>B</b>	170	12	14	28	-1,11	0,58	1,33	0,39	-0,13	0,55	3
<b>C</b>	155	11	12	29	-,84	0,58	-,90	0,39	-0,84	0	5
<b>D</b>	160	9	13	31	,35	0,58	-,38	0,39	0,05	0,68	2
<b>E</b>	165	8	16	32	1,37	0,58	,77	0,39	1,09	1	1

Tabela C. 2: Dados Simulados para A<sub>2</sub> (em unidades arbitrárias)

<b>Id</b>	<b>V<sub>1</sub></b>	<b>V<sub>2</sub></b>	<b>V<sub>3</sub></b>	<b>V<sub>4</sub></b>	<b>ECP<sub>2</sub></b>	<b>VC<sub>1</sub></b>	<b>ECP<sub>2</sub></b>	<b>VC<sub>2</sub></b>	<b>IBC</b>	<b>IBC<sub>std</sub></b>	<b>Order</b>
<b>A</b>	165	11	16	31	,15	0,59	,17	0,28	0,04	0,55	3
<b>B</b>	175	13	15	29	-1,34	0,59	1,02	0,28	-0,5	0	5
<b>C</b>	160	12	13	30	-,60	0,59	-1,22	0,28	-0,01	0,5	4
<b>D</b>	162	10	14	32	,56	0,59	-,83	0,28	0,1	0,61	2
<b>E</b>	170	9	17	33	1,23	0,59	,87	0,28	0,97	1	1

Após coletar os dados, realiza-se uma PCA para obter os *scores* dos Componentes Principais (IBC) para cada “observação” em ambos os anos.

O primeiro Componente Principal explica uma grande porção da variação nos dados, o segundo uma outra parte da proporção total.

Foram constatados, nesse exemplo, três Componentes Principais, porém apenas os dois primeiros apresentaram autovalores  $\geq 1$ .

Considerando que o teste de esfericidade para a Matriz formada pelas 4 colunas (V<sub>1</sub>, V<sub>2</sub>, V<sub>3</sub>, V<sub>4</sub>), em ambas as tabelas, apresentou um *p-valor*  $\leq 0,05$ , considerou-se que os dados atendem aos requisitos para prosseguir com análise.

Em seguida, os autovetores são calculados e analisados. Um procedimento genérico que ilustra o processo, utilizando a linguagem *R*, é representado, a seguir:

```
eigen_res_A_i <- eigen (mtz_numerica_A_i, symmetric = T, only.values = T)
```

Uma vez calculado, o resultado é salvo na variável `eigen_res_A_i`. Um procedimento genérico que ilustra o processo de análise PCA, utilizando a linguagem *R*, é representado, a seguir:

```
res. PCA_i <- psych::principal(mtz_numerica_A_i, nfactors = sum(eigen_res_A_i$values > 1), rotate = 'varimax', scores = TRUE, residuals = TRUE, oblique.scores = TRUE, method = "regression", use = "all.obs")
```

Neste exemplo, a matriz de correlação de Pearson foi calculada internamente pelo algoritmo acima (`scores = TRUE`), não sendo necessário padronizar inicialmente os valores dos dados.

Na sequência, o  $IBC_n$  é calculado para cada tabela. A fórmula genérica para calcular um  $IBC_n$ , ilustrativo, é:

$$IBC_n = \sum_{j=1}^m \sum_{i=1}^n VC_i ECP_{ij}$$

onde ECP trata-se do escore do Componente Principal rotacionado, VC representa a Variância Compartilhada de cada Componente Principal rotacionado,  $i$  é o número total de ECP, e  $j$  denota o número de casos ou observações. O processo de construção deste índice, ilustrativo, será explicado em detalhes a seguir.

Uma vez que, é mais intuitivo interpretar um valor entre 0 e 1 do que, por exemplo, entre -0,84 e 1,09, os valores de  $IBC_n$  foram padronizados para a escala [0,1], e salvos na variável  $IBC_n^{std}$ . A fórmula geral para a padronização Min-Max aplicada é:

$$IBC_n^{std} = \frac{IBC - IBC_{\min}}{IBC_{\max} - IBC_{\min}}$$

onde  $IBC_{\min}$  e  $IBC_{\max}$  são, respectivamente, os valores mínimo e máximo de  $IBC_n$  e  $IBC_n^{std}$  o resultado padronizado para a escala [0,1] de  $IBC_n$ .

Ao final, os resultados relativos à ordenação das "observações" em função de  $IBC_n^{std}$  são reordenados do maior para o menor valor e salvos na coluna "Order".

As colunas "IBC", " $IBC_{std}$ " e "Order" então podem ser utilizadas para estabelecer diversas comparações e interpretações, possíveis, entre as "observações", já que todas estão em suas respectivas escalas. A seguir, alguns exemplos são possíveis, como:

- **Comparação Direta:** é possível afirmar que, com um valor de  $IBC_{std}$  de 1, a "observação" "E" obteve o primeiro lugar em ambos os anos, indicando uma "evolução" relativa consistente ou que; com o menor valor de  $IBC_{std}$ , 0, a "observação" "C" apresentou a pior evolução relativa entre todas as "observações". Isso indica, com base nos critérios avaliados, que ela não evoluiu de forma adequada em comparação às demais.
- **Comparação Relativa:** uma "observação" com  $IBC_{std}$  de 1 indica a melhor "evolução" relativa entre todas as observações, enquanto um valor de 0 indica a pior "evolução" relativa.
- **Ordenação:** a "observação" "A" melhorou sua posição de 4º lugar no ano  $A_1$  para o 3º lugar no ano  $A_2$ , enquanto a "observação" "B" caiu de 3º para 5º.
- **Varição nos Escores:** é possível constatar que, embora algumas "observações" tenham apresentado IBC semelhantes, pequenas diferenças podem resultar em alterações na evolução.
- **Tendências ao Longo do Tempo:** é possível constatar que a "observação" "A" apresentou um aumento relativo em  $IBC_{std}$  de  $A_1$  para  $A_2$ , isso sugere que houve alguma evolução em relação às outras "observações".
- **Tendências Absolutas:** observando as evoluções, pode-se identificar algumas tendências consistentes, como uma "observação" consistentemente evoluindo bem "A" e "D" ou uma "observação" apresentando melhoria "C".
- **Distribuição do "Evolução":** a análise da coluna  $IBC_{std}$  permite avaliar a distribuição da "evolução" entre as "observações". Se muitas



“observações” estão agrupadas próximas a 1, isso sugere que muitas estão apresentando uma evolução acima da média. Se estão agrupadas próximo de 0, muitas estão apresentando uma evolução abaixo da média.

- **Identificação de *Outliers*:** “observações” que se desviam significativamente da média ou estão isoladas nos extremos (próximo de 0 ou 1) podem ser consideradas outliers e merecem uma análise mais detalhada.
- **Comparação entre Grupos:** se existirem subgrupos dentro das observações, os valores padronizados permitem comparações entre esses grupos. Por exemplo, se as “observações” “A” e “B” pertencem a um grupo e “C”, “D” e “E” a outro, pode-se avaliar qual grupo teve um desempenho médio melhor. Para realizar a comparação é possível utilizar os resultados das colunas “IBC” ou “IBC<sub>std</sub>”.
- **Conexão com outras Variáveis:** Pode-se correlacionar IBC<sub>std</sub> com outras variáveis para entender os fatores que influenciam o desempenho. Por exemplo, se V1-V4 são variáveis explicativas, pode-se investigar qual delas tem uma relação mais forte com IBC<sub>std</sub>.
- **Análise Multianual:** com os valores padronizados para múltiplos anos, é possível realizar análise de série temporal para identificar tendências, ciclos ou padrões recorrentes no desempenho das observações ao longo do tempo.


Em resumo, as colunas “IBC”, “IBC<sub>std</sub>” e “Order”, por serem padronizadas, fornecem uma base comum para comparação, facilitando a interpretação e análise dos dados, especialmente quando se deseja comparar, por exemplo, resultados absolutos ou relativos entre observações ou ao longo do tempo.


## Apêndice D - Detalhamento Estrutural das Tabelas de Dimensão e Fato empregadas na Construção do *Data Warehouse*

As tabelas de dimensão são representadas pelo padrão  $d[nome]$ , enquanto as tabelas de fato são representadas pelo padrão  $f[nome]$ .

Os tipos de relacionamentos são utilizados para definir como as tabelas são relacionadas entre si em um modelo de dados. Um exemplo de tipo de relacionamento é a seta direita com uma linha cheia, que indica um relacionamento ativo entre duas tabelas, onde uma tabela é a tabela "um", identificada pelo valor "1" e a outra é a tabela "muitos", representada pelo valor "\*". Isso significa que para cada registro na tabela "um", pode haver muitos registros correspondentes na tabela "muitos". Um outro exemplo de tipos de relacionamentos é a linha tracejada, que pode representar um relacionamento desativado entre duas tabelas ou uma conexão entre uma tabela calculada e suas tabelas de origem.

O símbolo de soma ( $\Sigma$ ) ao lado de um atributo em uma tabela no Power BI representa que a coluna é uma medida agregada. Isso significa que a coluna contém valores numéricos que podem ser somados para obter um valor agregado, como a soma de vendas ou a contagem de clientes.

Já o símbolo de soma  no canto inferior esquerdo de uma coluna em uma visualização de tabela ou matriz no Power BI indica que a coluna contém um valor numérico agregado, como a soma, média, máximo ou mínimo dos valores naquela coluna. Isso é útil para identificar rapidamente os valores agregados em uma visualização e distinguir as medidas das outras colunas que podem conter dados brutos.

Por fim, o ícone  no canto inferior esquerdo de uma coluna em uma visualização de tabela ou matriz no Power BI representa a presença de uma medida. Essa medida é criada por meio de uma fórmula DAX e pode ser usada para agregar e resumir dados em um modelo de dados. O ícone de "fx" é uma indicação visual útil para ajudar a identificar rapidamente as colunas relevantes em uma visualização e distinguir as medidas das outras colunas que podem conter dados brutos.

A visualização das Figura E. 1, Figura E. 2 e Figura E. 3 ilustram a aplicação do esquema floco de neve no DW, que compõem o *dashboard*.

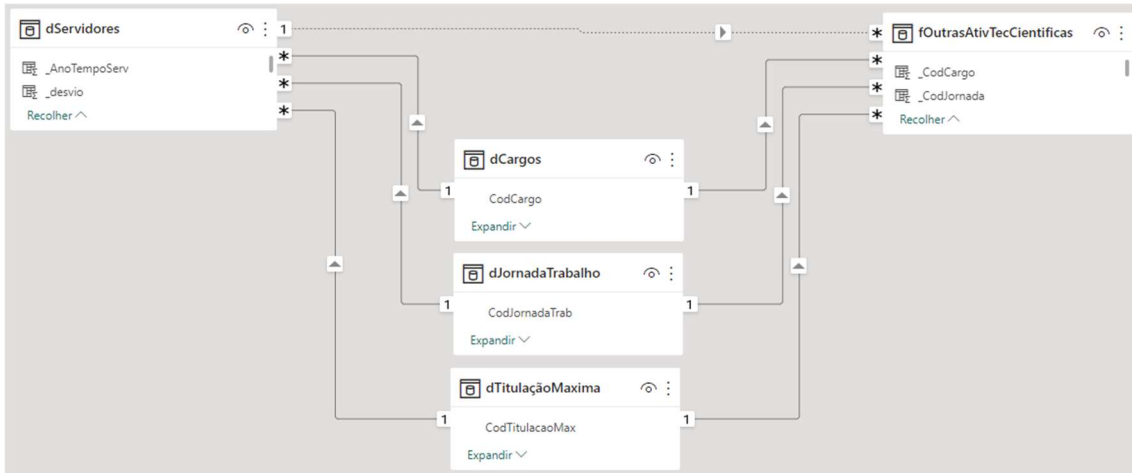


Figura E. 1: Representação do esquema floco de neve apresentando o relacionamento entre as tabelas dimensão e fato para a elaboração do dashboard - Módulo atividade PD&S: **Amplo**

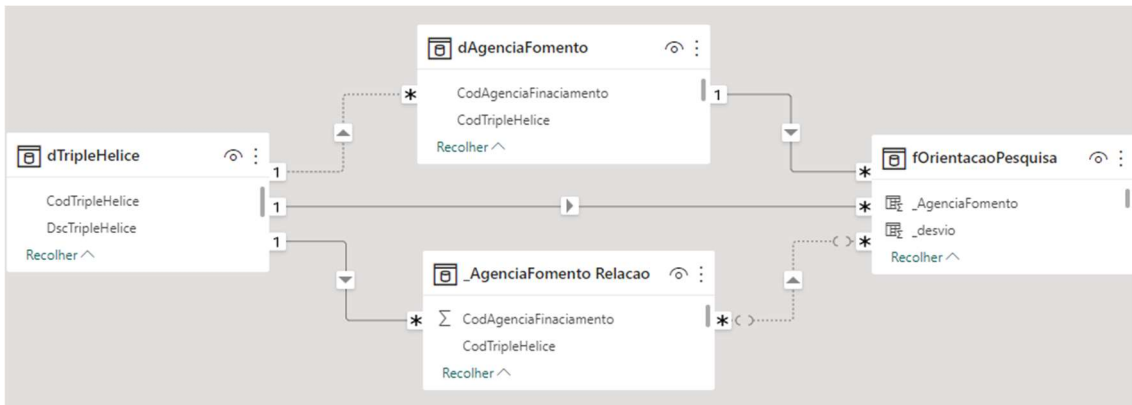


Figura E. 2: Representação do esquema floco de neve apresentando o relacionamento entre as tabelas dimensão e fato para a elaboração do dashboard - Módulo atividade PD&S: **Detalhado**

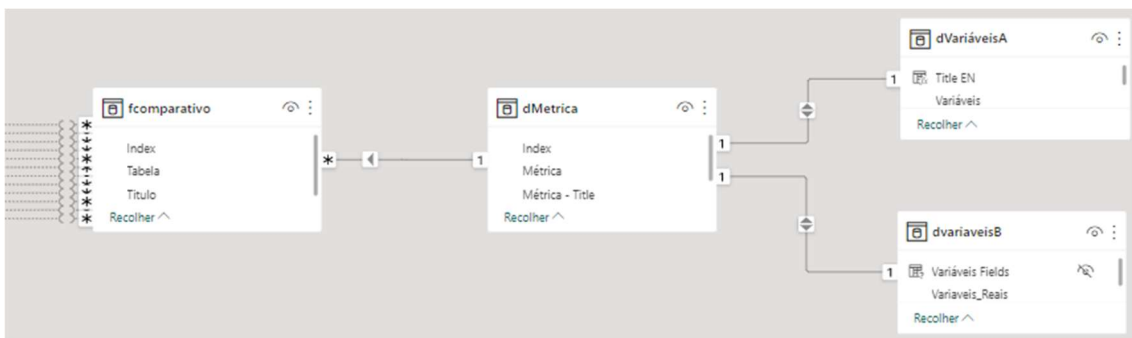


Figura E. 3: Representação do esquema floco de neve apresentando o relacionamento entre as tabelas dimensão e fato para a elaboração do dashboard - Módulo atividade PD&S: **Comparativo**

Nas tabelas dimensão e fato cada atributo métrico estará localizado em uma coluna (Hassan *et al.*, 2021) e, conseqüentemente, as observações em linhas, representando as ICTs.

## Apêndice E - Detalhamento do Processo de ETL

O Processo de Extração, Transformação e Carga (em inglês, *Extract, Transform, Load* - ETL) engloba técnicas dedicadas à coleta, estruturação e inserção de dados em um sistema de armazenamento ou *data Warehouse* ou *data Lake* (Ammar *et al.*, 2022; Diouf *et al.*, 2018; Inmon, 1995; Kimball *et al.*, 2008).

O processo de ETL é fundamental para preparar grandes volumes de dados para análise, garantindo que eles estejam corretos, consistentes e prontos para serem consultados e analisados. Seu detalhamento é essencial para compreender como os dados são coletados, limpos e carregados, estabelecendo assim a fundação para análises subsequentes e para a aplicação de técnicas como os Métodos de Interdependência Multivariada Exploratória. Para assegurar o sigilo comercial e proteger o segredo industrial, alguns detalhes metodológicos deste estudo foram preservados, sobretudo em face de possíveis negociações com investidores (Seção 3)

### E.1 - Raspagem dos Dados

A raspagem automatizada utilizou a linguagem *R* (Team, 2022) com o suporte da IDE RStudio como ambiente de desenvolvimento estatístico e de gráfico integrado baseado em software livre. A biblioteca do *R* utilizada foi a RSelenium – V 1.7.7. Os **Procedimentos** 1 ao 8 sintetizam o processo de extração adotados para cada caso, de acordo com cada base de dados considerada.

O **procedimento 1** extraiu os códigos identificadores das instituições da RFEPCT, conforme detalhado no Apêndice B, Quadro B. 13, F<sub>1</sub>.

O **procedimento 2** extraiu dados dos servidores públicos federais da RFEPCT, conforme Apêndice B, Quadro B. 13, F<sub>4</sub>.

O **procedimento 3** extraiu os números identificadores do Lattes dos servidores públicos federais ativos vinculados à RFEPCT, conforme detalhado no Apêndice B, Quadro B. 13, F<sub>2</sub>.

O **procedimento 4** foi usado para extrair dados das despesas federais em pesquisa e desenvolvimento vinculadas à RFEPCT das fontes listadas no Apêndice B, Quadro B. 13 (F<sub>1</sub>, F<sub>3</sub> e F<sub>7</sub>). Dos dados, 98,4% (equivalente a 462 observações de 42 ICTs ao longo de 11 anos) foram coletados do CES (F<sub>3</sub>). Apenas 0,6% (3 observações) foram

complementados com informações da LOA (F<sub>1</sub>) e RGIEP (F<sub>7</sub>). Mesmo após contatos com as 42 ICTs, todas indicaram a inexistência desses dados em suas bases, sugerindo a consulta nas fontes mencionadas.

O **procedimento 5** foi usado para extrair dados dos códigos Lattes dos servidores públicos federais da RFEPCT, utilizando o extrator Lattes da Plataforma Lattes (Apêndice B, Quadro B. 13, F<sub>5</sub>). O acesso ao extrator foi solicitado oficialmente pela Pró Reitoria de Pesquisa devido a demandas institucionais da época.

O **procedimento 6** foi utilizado para extrair dados no formato XML da Plataforma Lattes (Apêndice B, Quadro B. 13, F<sub>5</sub>). Rotinas internas foram desenvolvidas para criar *Data Marts* como pesquisa, publicações científicas e serviços.

No **procedimento 7**, utilizou o código único das ICTs para extrair dados do INPI (Apêndice B, Quadro B. 13, F<sub>6</sub>) sobre Desenvolvimento Tecnológico (Patente, Desenho Industrial, Programas de computador e Marcas) relacionados à RFEPCT. Esse método, diferente do padrão, evita problemas de autenticidade ao vincular registros às respectivas ICTs titulares.

O **procedimento 8** foi usado para extrair dados do RGIPE (Apêndice B, Quadro B. 13, F<sub>7</sub>), abrangendo resultados de PD&S, editais e ações de promoção da interação TH, vinculados à RFEPCT. Devido à falta de padronização e atualização dos documentos, os arquivos PDF extraídos foram salvos em uma pasta específica, servindo como informação complementar quando necessário.

Ao final de cada procedimento, matrizes de dados foram estruturadas em formato linha x coluna e armazenadas em *Data Marts* específicos no *Data Warehouse* (Subseção 3.5, Figura 2). Estas tabelas foram categorizadas conforme as características dos dados de cada fonte e o propósito deste trabalho. A Tabela D. 1 mostra um exemplo dos dados brutos de patentes extraídos usando a linguagem R.

Tabela D. 1: Exemplo da tabela patentes (INPI), extraída utilizando a linguagem R, contendo os dados brutos no formato linha x coluna utilizado para inserção no banco de dados temporário ao final do processo de ETL

ID	Cod IES	Cod CNPJ	Dsc link	Cod Pedido
1	2658	7510873000190	<a href="https://busca.inpi.gov.br/...">https://busca.inpi.gov.br/...</a>	BR 10 2021 0210435
2	2658	7510873000190	<a href="https://busca.inpi.gov.br/...">https://busca.inpi.gov.br/...</a>	BR 10 2021 0191058
3	2658	7510873000190	<a href="https://busca.inpi.gov.br/...">https://busca.inpi.gov.br/...</a>	BR 10 2021 0184014
4	2658	7510873000190	<a href="https://busca.inpi.gov.br/...">https://busca.inpi.gov.br/...</a>	BR 10 2021 0152082
5	2658	7510873000190	<a href="https://busca.inpi.gov.br/...">https://busca.inpi.gov.br/...</a>	BR 10 2021 0125063

## E.2 - Limpeza

No mapeamento, campos sem registros foram identificados como NULL ou NA. Embora não tenham sido encontrados dados NULL, os registros NA foram rotulados como “não informados”. Estes casos são comuns, por exemplo, no Lattes, onde a inserção manual dos dados e o não preenchimento de campos não obrigatórios resultam em registros incompletos. No CES, os dados de “Investimento em P&D” ( $X_1$ ) foram validados pelo Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP). Conforme sua metodologia, valores registrados como R\$ 0,00 indicam zero recursos e não NA.

## E.3 - Duplicação

Registros duplicados foram identificados e removidos utilizando a função `distinct()` do pacote `dplyr`, que mantém apenas a primeira ocorrência de valores únicos nas colunas selecionadas. Esta correção foi crucial ao analisar a quantidade de artigos por ICTs, já que um artigo pode envolver múltiplos atores de uma mesma ICTs. Critérios como correspondência entre autores e identificação de artigos em diferentes fontes foram aplicados para determinar duplicatas. Após a remoção automatizada, uma revisão manual foi realizada para detectar possíveis casos atípicos. Esta etapa assegurou a integridade dos dados e a precisão da análise subsequente.

#### E.4 – Revisão de Formato

Os dados foram padronizados para formatos consistentes, como a conversão de datas para somente o “ANO”, uma vez que era a informação comum em todas as fontes. Instituições com diferentes nomenclaturas, como “Instituto Federal de Alagoas”, “IFAL” e “Escola Técnica Fed de AL”, foram unificadas sob uma sigla única, “IFAL”. Essa padronização foi associada ao CNPJ reduzido, considerando somente os primeiros oito dígitos, devido à limitação das fontes de dados.

Na linguagem *R*, a função `tolower()` foi usada para padronização de caracteres e `gsub()` para substituições, como trocar espaços por traços. A biblioteca `lubridate()` auxiliou na manipulação de datas, com a função `year(ymd(data))` para extrair anos.

A natureza dos dados e o objetivo da análise determinaram a abordagem de padronização. Em algumas situações, foi necessário padronizar os dados utilizando a escala *Z-score* ou Min-Max, considerando-se o entendimento do objetivo deste estudo.

#### E.5 – Reestruturação de Chaves

Ao estabelecer chaves únicas, consultou-se a base de dados da Receita Federal para obter o CNPJ de cada ICTs, garantindo a correta identificação das instituições na base do INPI. Usando o CNPJ, os 8 primeiros dígitos representam a matriz e os subsequentes, as filiais. Isso nos permitiu discernir a matriz de cada ICTs e relacioná-la adequadamente nas tabelas. Desafios foram enfrentados quando apenas o CNPJ da matriz estava acessível. Isso levou à criação de um algoritmo para identificar filiais potenciais. À medida que os dados eram coletados, chaves únicas para interligar as tabelas eram geradas.

#### E.6 – Filtragem

Para selecionar as linhas e colunas pertinentes das fontes de dados (Apêndice B, Quadro B. 13), utilizou-se a linguagem *R* com funções da biblioteca `dplyr`, como `subset()`, `filter()`, `select()`, entre outras. No contexto das ICTs da RFEPCT, focou-se exclusivamente nos dados associados a essas instituições e aos canais de PD&S. Os dados relacionados as variáveis selecionadas para este estudo (Apêndice B, **Erro! Fonte de referência não encontrada.**) foram filtrados. Ao final do processo gerou-se as matrizes de dados

contendo as 42 ICTs da RFEPCT ou as 462 separadas anos a ano e suas respectivas variáveis selecionadas.

Durante a seleção, surgiram desafios, como a dispersão de informações vitais em diferentes fontes e inconsistências nos dados. A seguir, alguns exemplos identificados, como ICTs, que possuíam várias denominações diferentes, registros de atividades cadastrados como pesquisas quando na verdade deveriam pertencer a outras características (situação, identificada e questionada pelo especialista 8 – Apêndice B, Quadro B. 3); situações como essas e outras são frequentemente encontradas nas múltiplas fontes de dados oficiais utilizadas (Apêndice B, Quadro B. 13).

### **E.7 – Unindo Dados**

A integração de dados de várias fontes, incluindo informações das ICTs e coordenadas de longitude e latitude, possibilitou a visualização geoespacial da densidade de "patentes publicadas" ( $X_{20}$ ) por região em um mapa. Para integrar os dados, utilizou-se técnicas como mesclagem de tabelas para assegurar a precisão e coesão dos dados. Durante o processo, desafios foram enfrentados como, por exemplo dados faltantes e informações não correspondentes, necessitando aplicar técnicas de limpeza e padronização dos dados. Para esses procedimentos utilizou-se a linguagem *R*, aplicando as funções do pacote *dplyr*, como `left_join()`, para unir tabelas como “tabela\_coords”, que contém informações geográficas, e “tabela ICTs”, que detalha o número de patentes e publicações (Apêndice B, Figura B. 1).

### **E8 – Dividindo Dados**

A técnica de separação de dados em colunas distintas é essencial para assegurar a qualidade e coerência das informações geradas. Assim, os dados de uma coluna podem ser divididos em outras colunas ou tabelas. Neste exemplo, a tabela `TblPatentes` foi criada com as quatro colunas mencionadas como `CodIEP`, `CodCNPJ`, `CodPedido` e `PagHtml`. Em seguida, a função `separate_rows()` foi aplicada para separar os dados da coluna "CodPedido" em "*n*" linhas, utilizando o separador ";". Por fim, a função `rename()` foi utilizada para renomear a coluna "CodPedido" para "CodOrder". O resultado foi uma nova tabela chamada "tblPatentesPedido" contendo as colunas "CodIEP", "CodCNPJ" e "CodOrder", cujo dados foram separados linha a linha. Isso pode ser feito com o seguinte código em linguagem *R*:



```
tblPatentesPedido ← TblPatentes %>% separate_rows(CodPedido, sep = ";") %>%  
rename(CodOrder = CodPedido).
```

## E9 – Validando Dados

Para contorná-los, foi essencial a adoção de uma chave de controle que integra informações como SIAPE, CPF e input\_lotacao. Além disso, foi fundamental considerar que, ao mudar de ICTs, o servidor reinicia sua contabilização de contribuições para os canais de PD&S, evitando assim, a duplicação ou transposição incorreta de registros anteriores. Por exemplo, para validar os dados dos servidores, a função `inner_join()` do pacote `dplyr` foi aplicada, comparando as tabelas `tblServidores` e `tblAtuacaoProfissional`. As colunas "NomServidor", "CodSiglaOrgaoIE", "CodLATTES", "AnoIngressoOrgao" e "CodServidor" serviram como chaves de validação. A tabela resultante contém somente os registros que coincidem em ambas as tabelas. Por exemplo, entre 2008 e 2018, dos 77.789 servidores vinculados à RFEPC (42.215 professores e 35.574 técnicos), conseguiu-se recuperar o Currículo LATTES de 57.332 (38.638 professores e 18.694 técnicos). Desse grupo, apenas 41.360 tiveram pelo menos um registro de PD&S no período, dados esses que foram consolidados e carregados na ferramenta proposta. Isso pode ser feito com o seguinte código em linguagem *R*:

```
validacao ← inner_join(tblServidores, tblAtuacaoProfissional, by = c("NomServidor",  
"CodSiglaOrgaoIE", "CodLATTES", "AnoIngressoOrgao", "CodServidor")).
```

### E.10 – Sumarização

Nesta etapa foi fundamental representar os valores resumidos para obter seus totais, os quais foram calculados e armazenados em diferentes níveis como métricas ou indicadores de performance (também conhecidos em inglês, como KPIs). Por exemplo, é possível somar todos os investimentos em P&D que uma ICT recebeu para criar uma métrica de valor total dos investimentos públicos federais em P&D da ICT, denominada "ValTotalValInvestPD". Dessa forma, é possível avaliar a importância dos investimentos no fomento aos canais de PD&S das ICTs.

O uso de ferramentas computacionais, como a linguagem *R* pode auxiliar nesse processo por meio de funções como a `summarize()` do pacote `dplyr`, que permite a

agregação dos dados e o cálculo de métricas de negócios de forma eficiente e automatizada. Isso pode ser feito com o seguinte código em linguagem R:

```
ValTotalValInvestPD <- tblInvestimentosP&D %>% group_by(CodIEP) %>%  
summarise(ValorTotalInvestPD = sum(Investimento)).
```

### E.11 – Agregação

Durante a raspagem de dados, informações de diversas fontes foram consolidadas em tabelas separadas para análise posterior (Figura 2, Subseção 3.5 e Apêndice B, Figura B. 1). Esta abordagem facilitou a integração de dados de diferentes origens (Apêndice B, Quadro B. 13 e Figura B. 1), potencializando sua análise. Por exemplo, para combinar dados de "tblServidores" e "tblAtuacaoProfissional", considerou-se a agregação do número de servidores ativos por ano de ingresso na ICT e seu órgão vinculado. Para isso, pode-se utilizar a função `group_by( )` do pacote `dplyr`, que agrupa os dados por uma ou mais variáveis específicas, e a função `summarise( )` para realizar a contagem dos servidores por grupo. Isso pode ser feito com o seguinte código em linguagem R:

```
servidores_por_ano_orgao <- tblServidores %>% group_by(AnoIngressoOrgao,  
CodSiglaOrgaoIE) %>% summarise(NumServidores = n()).
```

### E.12 – Integração

Nesta etapa, cada métrica de dados única foi nomeada de forma padronizada após a integração, garantindo a consistência dos nomes e valores. Posteriormente, os dados foram armazenados temporariamente em um Modelo de Entidade de Relacionamento usando SQL (Apêndice B, Figura B. 1), seguindo o processo de ETL (Apêndice D e Seção 3.5.3). A normalização das tabelas foi realizada para prevenir redundâncias e inconsistências, com a implementação de chaves primárias e estrangeiras para assegurar a integridade dos dados. Finalmente, os dados foram transferidos para o Power BI (Seção 3.5, Figura 2), ferramenta escolhida para desenvolver a solução.

## Apêndice F - Detalhamento da Concepção do Projeto de Design UX para os Três Módulos (*Dashboard*) propostos para o MVP

O projeto conceitual do *UX design* proposto para o *dashboard* desenvolvido neste projeto foi dividido em três módulos. Cada módulo apresenta seu conjunto de filtros, métricas e KPIs que refletem as particularidades da(s) métrica(s) analisada(s) (Apêndice B, Quadro B. 14), a saber:

- a. **Amplio:** Neste módulo, é possível acessar informações gerais sobre as ICTs para as 14 macro variáveis consideradas ( $X_1, X_2, X_3, X_4, X_5, X_6, X_{10}, X_{11}, X_{12}, X_{13}, X_{18}, X_{19}, X_{20}, X_{21}$ ). Além disso, são apresentadas as respectivas métricas e KPIs, juntamente com possíveis (dis)similaridades entre as ICTs. Para cada macro variável, múltiplos filtros dinâmicos e interativos estão disponíveis, possibilitando a análise individual e precisa de cada uma.
- b. **Detalhado:** Neste módulo, cada uma das 14 macro variáveis consideradas ( $X_1, X_2, X_3, X_4, X_5, X_6, X_{10}, X_{11}, X_{12}, X_{13}, X_{18}, X_{19}, X_{20}, X_{21}$ ) é analisada com o maior detalhamento possível, permitindo obter informações mais específicas dos dados. Para esse módulo, optou-se por utilizar a variável “*Orientação de Pesquisa*” ( $X_3$ ) como exemplo para demonstrar sua funcionalidade. Além disso, nesse módulo **Detalhado**, os usuários têm acesso as mesmas funcionalidades do módulo **Amplio**. Adicionalmente, este módulo apresenta algumas funcionalidades específicas, como o gráfico de taxa de crescimento ano a ano (YoY%) e o filtro inteligente (em inglês, Q&A). O YoY% permite comparar dados em diferentes períodos e contextos, facilitando a identificação de tendências e padrões ao longo do tempo. A Q&A permite que o usuário faça perguntas sobre os dados que estão no sistema e ele lhe apresentará a resposta em formato de gráficos e tabelas. Por exemplo, você pode perguntar "Qual é o total de projetos de pesquisa por região, entre 2012 e 2015, que investigaram o tema vidro e qual a instituição com a maior taxa de crescimento YoY% nesse período?" e obter uma resposta visual imediata. Destaca-se que a razão pela qual a funcionalidade Q&A só está disponível na versão *offline* é que ela requer uma quantidade significativa de recursos de processamento e, portanto, pode ser mais lenta em uma conexão de internet menos potente; outra

situação é que as perguntas devem ser realizadas na língua inglesa, dentre outras questões. Dessa forma, o usuário pode explorar os dados de forma mais minuciosa e obter insights específicos sobre cada uma das macros variáveis. O *UX design* do *dashboard* foi planejado para facilitar a navegação e a interpretação das informações, permitindo uma análise mais profunda e abrangente dos dados.

- c. **Comparativo:** Neste módulo, são consideradas as 14 macro variáveis e suas interações com os **SPP** ( $X_1$  até  $X_{24}$ ), possibilitando a comparação de todas elas simultaneamente. No entanto, é recomendável ter cautela na decisão de comparar todas as variáveis ao mesmo tempo, devido à quantidade de informações geradas. Sugerimos que seja feita a comparação de até três variáveis por vez, para que a interpretação da informação seja mais objetiva e menos complexa. Diferente dos dois módulos **Amplio** e **Detalhado**, nesse optou-se, como exemplo de possibilidades implementar a análise multivariada dos dados utilizando a técnica correlação (Fávero *et al.*, 2020). Essa possibilidade ressalta que o MVP proposto pode incorporar outras técnicas e abordagens mais completas e robustas para analisar os dados e fornecer além de informação, o conhecimento. Para permitir que a análise de correlação possa ser realizada foi necessário padronizar os dados utilizando, por exemplo o padrão *z-scale*<sup>35</sup> (Fávero *et al.*, 2020). Essa ação foi possível uma vez que as métricas apresentam unidades distintas em unidades, dezenas centenas e até milhões.

---

<sup>35</sup> O z-score  $Z_i$  pode ser calculado subtraindo-se o valor individual da métrica (variável  $X_i$ ) da média  $\mu$  dessa mesma variável, dividindo sua diferença pelo desvio padrão  $\sigma$  dessa mesma variável:  $Z_i = (X_i - \mu) / \sigma$ .

A Figura F. 1 representa o esquema geral de *UX design* do projeto conceitual e da interface visual e de navegação do *dashboard*, proposto como MVP.

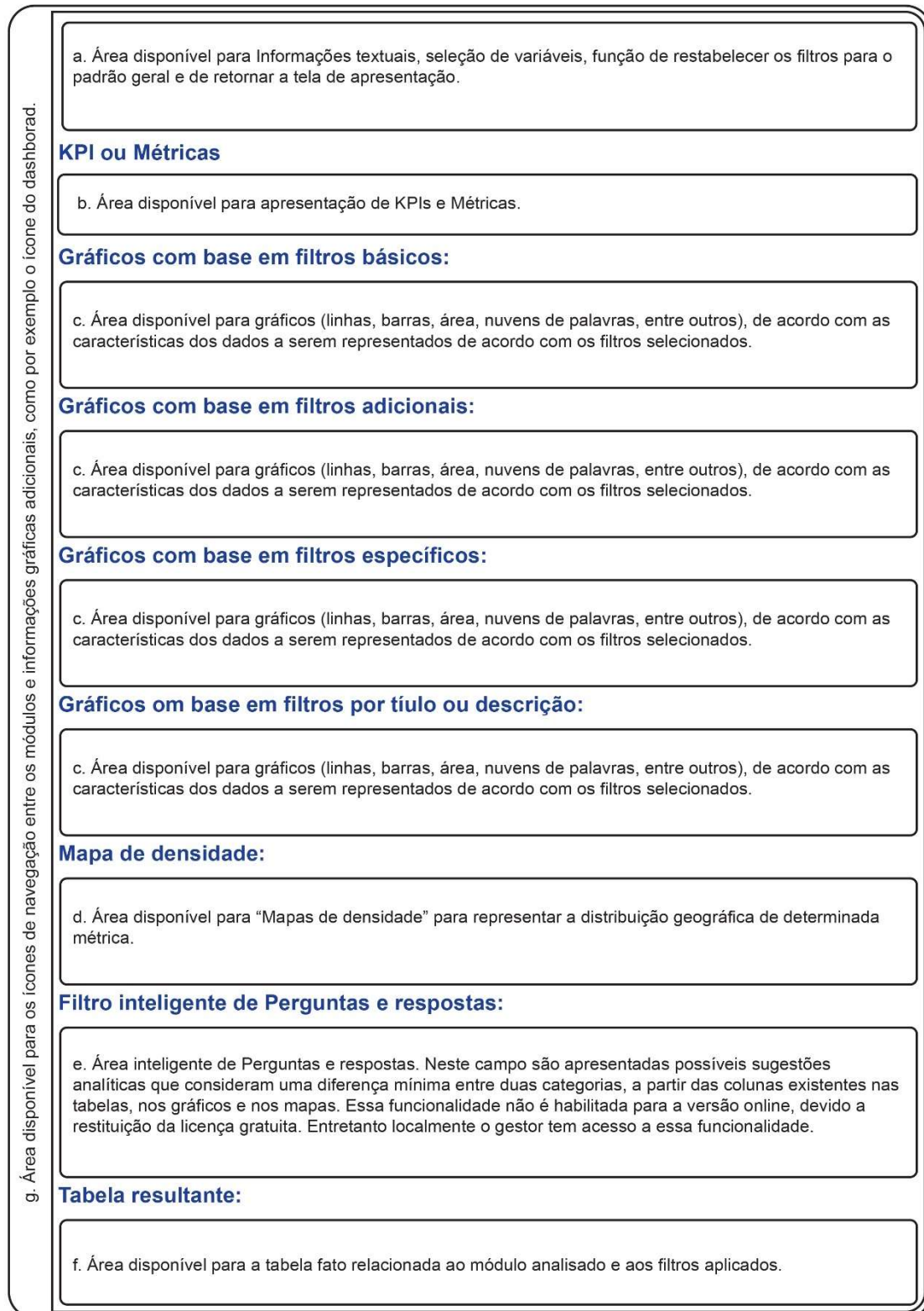


Figura F. 1:Esquema Geral de *UX design* do dashboard MVP

A estrutura básica do *UX design* foi planejada para agregar filtros e áreas adicionais que ampliem as possibilidades de interação e análise relacionadas às métricas e KPIs de forma dinâmica.

A Figura F. 1 apresenta o modelo básico de *UX design* proposto para o *dashboard*, que foi adaptado para cada um dos três módulos, conforme detalhado anteriormente nesta subseção. Dessa forma, a visualização da Figura F. 1 deve permitir ao usuário compreender como a estrutura básica do *dashboard* foi desenvolvida e como ela pode ser personalizada para as particularidades de cada módulo.

Os três módulos apresentam um conjunto de filtros disponíveis. Compreender a importância dos filtros a serem utilizados no *dashboard* proposto é fundamental para que o usuário possa explorar as múltiplas possibilidades de análise oferecidas pelo MVP.

Os filtros dinâmicos e interativos foram desenvolvidos para permitir uma exploração mais detalhada dos dados e a obtenção de informações relevantes de forma rápida e objetiva.

Os filtros disponíveis no *dashboard* são diversos, incluindo seleção de período, região, área temática, tipo de projeto, tipo de financiamento, entre outros. É possível combinar cada filtro com os demais, o que permite uma análise mais completa e a identificação de tendências e padrões específicos nos dados. Esses filtros foram definidos previamente durante o processo de modelagem dimensional para aprimorar a qualidade e eficácia das análises realizadas no *dashboard*.

A Figura B. 14 (Subseção 3.5.2), apresenta as dimensões "Região", "UF", "Sigla", "Instituição" e "Ano" que correspondem às colunas de 1 a 5, respectivamente. Esses foram definidos como filtros padrões, por terem sido aplicados a todos os 3 módulos do *dashboard*. De forma geral, os filtros fornecem informações mais precisas e detalhadas sobre as métricas de cada módulo, permitindo a análise de dados em diferentes perspectivas, inclusive possibilitando a realização de filtros múltiplos.

Os filtros de pesquisa na interface subdividem-se em três categorias:

- a. Gerais, que se propagam para todos os módulos;
- b. Adicionais, que são aplicados apenas a determinadas atividades;
- c. Específicos, que são aplicados às métricas, considerando seus microdados disponíveis.

A interface genérica do *dashboard* subdivide-se em:

- a. Informações textuais, seleção de variáveis, função de restabelecer os filtros para o padrão geral e de retornar a tela de apresentação;

- b. Área para apresentação de KPIs e Métricas;
- c. Áreas para gráficos (linhas, barras, área, nuvens de palavras, entre outros), de acordo com as características dos dados a serem representados de acordo com os filtros selecionados;
- d. Área para “Mapas de Densidade” para representar a distribuição geográfica de determinada métrica;
- e. Área inteligente de Perguntas e Respostas. Neste campo são apresentadas possíveis sugestões analíticas que consideram uma diferença mínima entre duas categorias, a partir das colunas existentes nas tabelas, nos gráficos e nos mapas. Essa funcionalidade não é habilitada para a versão *online*, devido a restituição da licença gratuita. Entretanto localmente o gestor tem acesso a essa funcionalidade;
- f. Área disponível para a tabela fato relacionada ao módulo analisado e aos filtros aplicados;
- g. Área disponível para os ícones de navegação entre os módulos e informações gráficas adicionais, como por exemplo o ícone do *dashboard*.

Além disso, essa interface apresenta as seguintes características:

- a. Os gráficos, tabelas e mapas são interativos e interconectados considerando os filtros aplicados;
- b. A função Mapa é utilizada para apresentadas como a densidade dos dados de uma determinada métrica é distribuído geograficamente.

Esta subseção apresentou a concepção e estrutura básica dos painéis visuais (*dashboards*) desenvolvidos na plataforma Power BI, com um *UX design* que agrega filtros e áreas adicionais para ampliar as possibilidades de interação e análise de métricas e KPIs de forma dinâmica. O projeto conceitual do *dashboard* foi dividido em três módulos, cada um apresentando seu conjunto de filtros, métricas e KPIs refletindo suas funcionalidades. A importância dos filtros disponíveis para explorar as múltiplas possibilidades de análise oferecidas pelo MVP foi destacada, assim como a Figura F. 1 (Apêndice F), que apresenta o esquema Geral de *UX design* do *dashboard*, representando, por exemplo a localização visual dos filtros definidos no processo de modelagem dimensional.

## Apêndice G - Detalhamento do Processo de Validação do Projeto de *UX Design* para os Três Módulos (*Dashboard*) propostos para o MVP

O MVP proposto foi validado por meio de estudos de caso e por parecer técnico emitido por especialistas vinculados ao ecossistema das ICTs, SPP e GOV. A validação teve como objetivo verificar a efetividade da solução proposta e identificar possíveis ajustes ou melhorias necessárias a serem implementadas em trabalhos futuros.

Existem várias metodologias que podem ser utilizadas para validar soluções que envolvam o desenvolvimento de produto, processo ou serviço por especialistas. A escolha da mais adequada depende do contexto e dos objetivos da validação. Algumas metodologias que podem ser consideradas são:

- Validação por especialistas (Nielsen, 1994; Rozmi *et al.*, 2021): a validação consiste em convidar especialistas para avaliar produto, processo ou serviço, considerando adequação aos objetivos, clareza das visualizações, facilidade de uso e qualidade dos dados, entre outros aspectos relevantes. As revisões podem ser individuais ou em grupo. O feedback dos especialistas pode ser coletado por meio de relatórios técnicos, entrevistas, questionários ou observação direta (Apêndice I). Os resultados ajudam a não só a identificar problemas a serem corrigidos, bem como identificar se a proposta atende aos parâmetros técnicos e/ou conceituais inicialmente estabelecidos. Esse processo pode conter várias etapas até que o produto atinja seu maior grau de maturidade tecnologia.
- Análise heurística (Nielsen, 1994): consiste em avaliar um produto, processo ou serviço com base em um conjunto de heurísticas pré-definidas de usabilidade. Essas heurísticas são critérios de avaliação baseados em princípios de *design* e usabilidade e podem incluir critérios como facilidade de navegação, clareza de informações, consistência, entre outros. A análise heurística pode ser realizada por especialistas em usabilidade e *design*.
- Análise de dados (Few, 2006): consiste em avaliar a precisão e a validade das informações apresentadas em ambientes que utilizam dados como produto principal. Esse tipo de validação pode ser feito por especialistas em análise de dados ou estatística, que podem avaliar a qualidade dos



dados, a consistência dos resultados e a precisão das conclusões apresentadas.

Em geral, a literatura recomenda que a combinação de várias dessas metodologias pode fornecer uma avaliação mais completa e confiável de um produto, processo e/ou serviço. Outro ponto a destacar é a importância de documentar todas as etapas da validação e garantir que as modificações feitas sejam baseadas nos resultados da validação. As etapas de validação consideradas foram:

#### **a) Objetivo da Validação**

A validação tem como objetivo garantir que o MVP proposto atenda às necessidades e expectativas do público-alvo, que são pesquisadores, gestores de ICTs e tomadores de decisão. A validação da ferramenta deve considerar aspectos como:

- **Relevância:** refere-se à importância e pertinência da ferramenta em relação às necessidades e objetivos do usuário, ou seja, se a ferramenta é útil e aplicável para o público-alvo.
- **Capacidade de contribuição:** diz respeito ao valor agregado que a ferramenta oferece, ou seja, se ela contribui de forma significativa para os processos e atividades do usuário.
- **Objetividade das informações apresentadas:** refere-se à clareza, imparcialidade e precisão das informações fornecidas pela ferramenta, sem viés ou interpretação subjetiva.
- **Capacidade de atender às demandas e expectativas do público-alvo:** diz respeito à capacidade da ferramenta de atender às necessidades e expectativas do usuário, em relação a funcionalidades, interface, usabilidade, entre outros aspectos.
- **Facilidade de uso e compreensão:** refere-se à facilidade e clareza de uso da ferramenta, incluindo a facilidade de navegação, a compreensão das informações apresentadas e a facilidade de interpretação.
- **Efetividade da visualização das informações:** diz respeito à capacidade da ferramenta de apresentar informações de forma clara e objetiva, permitindo uma análise precisa e confiável.

- Capacidade de fornecer *insights* relevantes a necessidade do usuário: refere-se à capacidade da ferramenta de fornecer informações relevantes e úteis para o usuário, permitindo a identificação de tendências, padrões e oportunidades de melhoria.
- Capacidade de contribuir para tomada de decisões baseadas em informações confiáveis e objetivas: diz respeito à capacidade da ferramenta de contribuir para a elaboração de estratégias, políticas públicas e tomadas de decisão com base em informações confiáveis e objetivas, evitando prejuízos econômicos, tecnológicos e científicos.
- Capacidade de eliminar os métodos subjetivos comumente empregados para tomadas de decisões relativas aos canais PD&S: Refere-se à capacidade da ferramenta de eliminar ou reduzir a subjetividade e o viés nas tomadas de decisão, fornecendo informações objetivas e confiáveis.
- Capacidade de contribuir para a resolução de problemas e para o suporte à tomada de decisão: diz respeito à capacidade da ferramenta de fornecer informações e insights que ajudem a resolver problemas e a tomar decisões mais informadas e eficazes.
- Capacidade de oferecer subsídios tecnológicos *online*: Refere-se à disponibilidade da ferramenta de forma online, permitindo o acesso remoto a informações e dados.
- Capacidade de análise objetiva: Refere-se à análise baseada em fatos e dados, sem viés ou interpretação subjetiva.
- Capacidade de análise confiável: Refere-se à análise baseada em informações confiáveis e precisas, com baixo risco de erros ou distorções.

#### **b) Seleção da Metodologia de Validação**

A literatura recomenda que a combinação de várias metodologias (consultar subseção 3.9.1) pode fornecer uma avaliação mais completa e confiável. Isso posto, para a avaliação do MVP proposto, optou-se pela metodologia baseada na revisão por especialistas sugerida por Nielsen (1994, p. 103), corroborada por Rozmi *et al.* (2021) e a análise de dados sugerida por Few (2006). Eles sugerem, em síntese, que:

- A utilização de um grupo de indivíduos especialista em determinada área para fornecer *feedback* sobre a importância, adequação e precisão do conteúdo de um instrumento de pesquisa ou modelo é um método válido e confiável;
- O processo de validação do instrumento deve garantir que os especialistas tenham acesso as mesmas orientações de como a avaliação deve ser conduzida;
- Os especialistas devem representar a diversidade do público-alvo que o modelo (MVP proposto) busca atender;
- Os especialistas tenham experiência de trabalho e preferencialmente tenham ou estejam exercendo cargos importantes em uma instituição na área a ser avaliada;
- O número apropriado de especialistas, que pode variar de três a cinco;
- Os resultados da validação devem ser registrados em um relatório detalhado e completo para demonstrar a confiabilidade e validade do instrumento de pesquisa ou modelo validado.

No Apêndice I, está disponível o modelo do documento de avaliação enviado por e-mail aos especialistas. Esse documento de avaliação levou em consideração alguns aspectos mencionados anteriormente, com ênfase naqueles que poderiam contribuir para responder ao objetivo e às hipóteses apresentadas neste trabalho. Ao envolver especialistas no processo de avaliação, buscou-se obter feedback valioso e insights sobre a relevância e a adequação das métricas, KPIs e gráficos selecionados, bem como a validade e a robustez das análises realizadas.

A inclusão de especialistas no processo de avaliação ajuda a garantir que as abordagens adotadas e os resultados obtidos sejam fundamentados na prática e tenham relevância no contexto das ICTs. Além disso, a colaboração com especialistas pode enriquecer o estudo, fornecendo uma perspectiva mais ampla e aprofundada sobre as questões em análise. Essa abordagem colaborativa é crucial para alcançar os objetivos estabelecidos e responder às hipóteses apresentadas de maneira eficaz e precisa.

### **c) Identificação dos Participantes**

De acordo com a literatura, deve-se especificar as características relevantes dos participantes, como conhecimento prévio da área, experiência com dashboards, entre outros. Assim sendo, a seleção dos especialistas para a validação envolveu pró-reitores de pesquisa e/ou inovação do ecossistema das ICTs, diretores de engenharia de dados do

ecossistema das SPP e diretores de fundação de amparo à pesquisa do ecossistema do GOV.

#### **d) Planejamento das Atividades**

Para que o especialista pudesse emitir seu parecer após o envio do material foi disponibilizado um prazo de 30 dias para enviar o retorno do documento preenchido. Após a recepção do documento foi coletado o *feedback* a ser analisado.

#### **e) Análise dos Resultados**

Nessa etapa, analisou-se os resultados da validação, identificando as contribuições e os pontos positivos. As sugestões de melhorias e problemas foram divididas em dois casos de aplicabilidade: i) imediata: realizada ainda neste trabalho, e futura: pós conclusão deste trabalho. Os resultados foram comparados com os critérios de sucesso definidos na etapa de descrição do objetivo. Na seção 4 os resultados foram apresentados de forma clara, evidenciando-se os pontos fortes e fracos destacado pelos especialistas em relação ao MVP proposto.

## Apêndice H - Simulação de Acesso ao MVP: Resultados do Usuário W de uma ICT

Ao acessar o MVP disponível no *link*: <https://prequest.websiteseguro.com/bi>, foi idealizado um *usuário W* que estaria interessado(a) em explorar os dados de sua IEPC visando entregar um relatório demandado pelo TCU referente aos seus resultados relacionados aos canais de PD&S entre 2008 e 2018. De certa forma, os entrevistados tiveram acesso ao MVP também enquanto usuários(as).

Para ter acesso aos dados e realizar as análises necessárias para obter as informações demandas com agilidade e precisão ele iniciou acessando o módulo **Amplio** desta ferramenta tecnológica para obter inicialmente uma compreensão global do que lhe foi solicitado.

Curioso em melhor entender a atividade *Projeto de P&D*, tal *usuário W* decidiu por clicar no ícone apropriado para ativar o módulo **Amplio**, seguindo as orientações fornecidas na Figura 4a – ícone superior (Subseção 4.5, Figura 2) , visualizará a Figura B. 15 (Apêndice B).

Ao ativar o módulo, o *usuário W* foi imediatamente apresentado a uma série de visualizações de dados relacionadas à atividade de *Projeto de P&D*, conforme ilustrado acima. Explorando as informações disponíveis, o *usuário W* começou a analisar os resultados apresentados em diferentes formatos, como gráficos de barras, nuvens de palavras, mapas de densidade e tabelas.

Nesta interface, os filtros básicos disponíveis são: ano, região, UF e ICTs. Os itens adicionais, quando aplicáveis, incluem: regime de trabalho, cargo, titulação máxima dos pesquisadores e intervalo de anos em que os pesquisadores atuam na RFEPCT. Os filtros específicos disponíveis são: grande área de conhecimento, situação, potencial de aplicação aos setores econômicos (CNAE) e título ou descrição.

Ainda no módulo **Amplio**, após os filtros determinados, o usuário teve acesso a um mapa de densidade e uma a tabela fato contendo todas as 28 métricas disponíveis, considerando o quantitativo dos dados relacionados as seleções e interações realizadas. Nesta simulação o(a) interessado(a) pode visualizar, por exemplo as regiões com maior ou menor densidade em relação ao número ou valor total da métrica selecionada após a aplicação dos filtros.

Na sequência, de acordo com as seleções e filtros realizados ele foi acompanhando os indicadores KPIs relacionados ao Tempo que incluem o tempo médio para conclusão

de Projetos de P&D (concluído ou em andamento, em anos); ao Custo que podem incluir, por exemplo, o custo médio dos projetos de P&D conduzidos por pesquisadores que receberam bolsas de pesquisa, durante o período de concessão da bolsa (um exemplo, que pode ser adicionado por solicitação do *usuário W*); as Métricas que incluem o número de Projetos de P&D por CNAE; quantidade de consultoria por titulação máxima do pesquisador; quantidade de patente publicada em parceria com os setores ICTs-SPP-GOV.

É importante destacar que nesta versão do MVP, os KPIs são apresentados de acordo com o contexto e configurados para representar os indicadores adequados para cada métrica analisada. Algumas das métricas disponíveis nesta plataforma incluem a quantidade de *projetos de P&D* por CNAE, a quantidade de consultorias por titulação máxima do pesquisador e a quantidade de patentes publicadas em parceria com os setores da tripla hélice, entre outras.

Como o *UX design* foi planejado para ser simples, objetivo e adaptado a diversos meios de acesso (por exemplo, *tablet*, *smartphone*, *notebooks* e *smart TVs*) de modo tal que o usuário *W* pode, se precisar se ausentar da ICT, continuar sua análise utilizando seu *smartphone* até ter acesso a um *notebook* ou computador de mesa para ter maior espaço de tela e assim ter mais conforto visual.

Ainda acessando o módulo **Amplo**, ele(a) pode decidir por comparar seus resultados gerais em relação à RFEPCT ou em relação às suas respectivas unidades. Com essa ação a ferramenta pode possibilitar a visualização dos resultados relacionados as métricas  $X_1:X_{24}$  e os KPIs de cada um dos canais PD&S disponibilizados nesse MVP para efeito comparativo.

As comparações podem ser realizadas de tal modo a selecionar diferentes ICTs ou aplicando filtros. A seguir, são apresentados alguns exemplos de KPIs relacionados ao tempo, custo e quantidade, e métricas possíveis de serem apresentadas a partir dos dados coletados.

Cada visualização de dados pode oferecer ao *usuário W* uma nova perspectiva sobre a atividade *Projeto de P&D*, permitindo-lhe mergulhar em diferentes aspectos e nuances dessa métrica. À medida que explora tais informações, o *usuário W* pode começar a identificar padrões e tendências, o que pode ajudar a entender melhor o cenário geral do projeto de pesquisa e desenvolvimento.

Ao vivenciar essa experiência, o *usuário W* pode perceber o potencial do módulo **Amplo** e em como poderia ser empregado para obter uma visão abrangente e informativa

sobre a atividade do *Projeto de P&D*. Essa experiência pode proporcionar ao *usuário W* uma base sólida para aprimorar seu conhecimento e tomar decisões mais fundamentadas no futuro. Decidido a aprofundar sua investigação, *W* pode buscar conhecer mais detalhadamente as características e informações relacionadas às orientações de pesquisa acessando o módulo **Detalhado**, uma vez que essas são a origem da maioria dos resultados associados aos canais de PD&S.

Após obter informações gerais sobre a situação de sua ICT, o *usuário W* pode decidir por aprofundar a investigação, optando por acessar o módulo **Detalhado**. Este módulo, diferentemente do geral, apresenta uma composição mais particularizada dos dados e disponibiliza gráficos, KPIs, métricas, mapas e tabelas mais detalhadas.

Ao clicar no ícone correspondente na Figura 4a – ícone intermediário (Subseção 4.3, Figura 4) o *usuário W* pode visualizar a Figura B. 16 (Apêndice B). Ao acessar esse módulo, o *usuário W* pode selecionar a métrica *Orientações de Pesquisa*. Os resultados do módulo **Detalhado** são expressos visualmente de forma semelhante ao módulo **Amplo**, utilizando gráficos, KPIs, métricas, mapas e tabelas, adicionado da área para pergunta e resposta inteligente (consultar, Apêndice F, Figura F. 1: item “b”).

Os filtros básicos e adicionais são semelhantes ao módulo **Amplo**, enquanto os específicos variam conforme a métrica selecionada, representando a menor granularidade dos dados. No exemplo proposto, filtros específicos abrangem situação, grande área de conhecimento, aplicação aos setores econômicos (CNAE), setor TH, relevância sugerida pelo pesquisador, bolsa, tipo/natureza, título e agência de financiamento. Algumas métricas e KPIs específicos incluem: região, UF, ICTs, ano, número de pesquisadores e orientações de pesquisa, e taxa média de orientação anual.

Na sequência, após os filtros e procedimentos sugeridos acima, tal usuário pode visualizar um mapa de densidade e uma a tabela fato contendo todas as 28 métricas disponíveis, similares ao modulo amplo. Nesta simulação tal usuário pode decidir por acompanhar de forma mais detalhada alguns indicadores relacionados a dimensão orientações de pesquisa. Os KPIs relacionados ao tempo incluem tempo médio para conclusão de orientações de pesquisa, enquanto os relacionados à qualidade incluem número de orientações de pesquisa (concluídas); ou ainda, por número de orientações de pesquisa (concluídas) por pesquisador, número de orientações de pesquisa (concluídas) em interação SPP por Pesquisador; Taxa média de crescimento de Orientações de Pesquisa ano a ano (média YOY%), dentre outras.

Dentre as Métricas , o usuário pode ter acesso a taxa média de crescimento de Orientações de Pesquisa ano a ano (média YOY%); ou ao Número de Orientações de Pesquisa aplicáveis aos Setores Econômicos vinculados ao CNAE (Quadro B. 16, Apêndice B) submetidas por cargo, jornada de trabalho, titulação máxima e intervalo de tempo de serviço; número de orientações de pesquisa vinculadas por setor tripla hélice submetidas por cargo, jornada de trabalho, titulação máxima e intervalo de tempo de serviço; número de orientações de pesquisa financiadas pelo SPP na modalidade bolsa.

Assim como no módulo **Amplo**, no **Detalhado** também é possível filtrar dinamicamente os dados por busca textual (título ou descrição) relacionada à métrica selecionada. O módulo **Detalhado** pode auxiliar na tomada de decisão, monitoramento e identificação de ações corretivas, permitindo análises mais detalhadas de cada métrica e aumentando a precisão da análise baseada em dados.

Através de gráficos, KPIs, métricas e mapas, bem como da área de pergunta e resposta inteligente, o *usuário W* pôde ainda explorar informações detalhadas e identificar padrões e tendências importantes. Com filtros específicos e interações adaptáveis, o *usuário W* pode ter uma compreensão mais completa das *Orientações de Pesquisa* e a possibilidade de analisar possíveis relações com os resultados de *Projetos de P&D* obtidos da análise do módulo **Amplo**, melhorando sua capacidade de tomar decisões informadas e monitorar ações corretivas, como também preencher as informações solicitadas no relatório do TCU de forma transparente, rápida e precisa.

Após analisar a situação da sua ICT, o *usuário W* pode comparar os resultados obtidos com os da sua região e o panorama nacional. Esta análise pode ser feita com base nos Projetos de P&D identificados no módulo **Amplo**. A abordagem detalhada pode permitir identificar padrões e tendências significativas ao considerar a métrica Orientações de Pesquisa no módulo **Detalhado**. Conseqüentemente, percebe-se a necessidade de explorar informações mais complexas relativas a essas métricas, pois tanto o usuário quanto sua equipe, além de outros membros que possam compor a IEPC poderiam estar convictos que essas métricas Projetos de P&D e Orientações de Pesquisa estão diretamente associadas, porém essa convicção não estaria baseada em dados.

Para elucidar o problema e dirimir as especulações, o *usuário W* pode optar por acessar o módulo **Comparativo**. Essa ação permitiria utilizar de técnicas estatísticas como correlações, item disponível na ferramenta para avaliar a relação linear entre duas ou mais variáveis quantitativas, permitindo uma avaliação precisa e embasada para a tomada de suas decisões. Esse processo pode ser crucial para o monitoramento de ações



corretivas no contexto do Projeto de P&D, o qual provavelmente avaliará. Tal procedimento também pode responder, com transparência e precisão, ao formulário do TCU, que havia solicitado informações sobre essa métrica em particular.

Ao clicar no ícone correspondente na Figura 4a – ícone inferior (Subseção 4.3, Figura 4) o *usuário W* pode visualizar a Figura B. 18 (Apêndice B). Ao acessar o módulo, o *usuário W* pode selecionar a métrica *Orientações de Pesquisa e Projetos de P&D*, conforme ilustrado nas Figura B. 17: Ilustração do *UX design dashboard* para o módulo – **Comparativo**. Os resultados do módulo **Comparativo** são expressos visualmente de forma semelhante aos módulos **Amplio** e **Detalhado**, adicionando-se uma matriz de correlação.

Deste modo, para investigar o problema em questão o usuário selecionou as métricas Projetos de P&D e Orientações de Pesquisa para analisar o grau de correlação e significância. Ao efetuar a seleção dessas duas métricas, automaticamente seus valores são padronizados aplicando-se o método de *Z-score*.

Os resultados expressos no gráfico de linhas se adaptam dinamicamente aos filtros aplicados. Na versão MVP, métricas ou KPIs relacionadas ao Custo, Tempo e Qualidade não foram disponibilizadas, mas podem ser adicionadas em versões futuras da plataforma.

Após analisar os resultados obtidos no módulo **Comparativo**, o *usuário W* pode observar, portanto, correlações estatisticamente significativas entre a atividade do *Projeto de P&D* e várias outras métricas. Por exemplo, houve uma forte correlação positiva entre *Projeto de P&D* e Pesquisadores  $r(42) = 0,94, p < 0,001$  e entre *Projeto de P&D* e *Extensão*  $r(42) = 0,93, p < 0,001$ . Tais correlações indicam que um aumento nas atividades do *Projeto de P&D* está fortemente relacionado ao aumento de pesquisadores e atividades de extensão acadêmica.

Também foram encontradas correlações positivas significativas entre *Projeto de P&D* e métricas como *Projetos de Pesquisa*, *Orientações de Pesquisa*, *Publicações* e outras atividades, o que demonstra a relação entre o Projeto de P&D e várias áreas do ecossistema de pesquisa e desenvolvimento.

Com base nas análises realizadas e nas informações obtidas por meio da ferramenta tecnológica fruto desta tese, o *usuário W* pôde compreender melhor as correlações e tendências entre as métricas, fornecendo informações valiosas para aprimorar suas estratégias e ações em ecossistema de pesquisa e desenvolvimento da RFEPCT. Assim, a ferramenta possibilitou ao *usuário W* realizar uma avaliação mais precisa e embasada em dados para a tomada de decisões e monitoramento de ações

corretivas no contexto do *Projeto de P&D* e responder ao relatório do TCU com transparência e objetividade, justificando suas decisões como gestor em fomentar determinadas medidas em prol de outras, considerando o projeto de desenvolvimento Institucional planejado.

## Apêndice I - Modelo de Documento encaminhado para os Avaliadores

### Assunto do e-mail:

Convite para análise de uma ferramenta em BI para canais de Pesquisa, Desenvolvimento Tecnológico e Serviços para o ecossistema das Instituições de ensino, pesquisa, ciência e tecnologia.

### Corpo do e-mail:

Prezado [nome],

Sou Cleber Nauber dos Santos, doutorando em Engenharia Industrial na UFBA, em fase de defesa. Gostaria de convidá-lo/a para participar da análise de um produto tecnológico, gerado durante o meu doutorado, sob a forma de um MVP (*Minimum Viable Product*) online, cujo objetivo da ferramenta é *"fornecer uma análise objetiva e confiável dos canais de Pesquisa, Desenvolvimento e Serviços do ecossistema das Instituições de Educação Profissional, Científica e Tecnológica (IEPCT/ICTs), com base em múltiplas fontes de dados, para embasar a tomada de decisão, análise, gestão e monitoramento. Essa análise deve estar relacionada às demandas do setor produtivo e do governo ao ecossistema das ICTs, e ser embasada em evidências confiáveis."*

O objetivo da tese é: *desenvolver uma solução tecnológica que possibilite a análise objetiva e confiável dos canais de PD&S em ecossistema das ICTs, utilizando dados obtidos a partir de múltiplas fontes, a fim de fundamentar a tomada de decisão, análise, gestão e monitoramento das demandas do setor produtivo e do governo.*

Para a validação da ferramenta utilizou-se os dados correspondentes à primeira década de atividade da Rede Federal de Educação Profissional Científica e Tecnológica (RFEPCT). Após a análise do estado da arte e da técnica revisada o tema, defendemos a tese de que a solução proposta como ferramenta online gera informações que podem contribuir com os gestores em suas tomadas de decisões, assim como dar suporte a servidores, pesquisadores, discentes e a comunidade em geral em seus mais diversos tipos de pesquisas.

A ferramenta é um sistema de *Business Intelligence* (BI), que utiliza dados de pesquisa, desenvolvimento e prestação de serviços científicos e/ou tecnológicos coletados de 42 ICTs vinculadas à RFEPCT. Ela permite análises para além dos modelos tradicionais existentes como os educacionais e de gestão como já ocorre na Plataforma Sucupira, apresenta um conjunto de informações relevantes para o ecossistema de C,T&I colaborando para o fomento das interações TH e foi planejada para integrar-se há outras soluções que venham contribuir para o ecossistema.

Gostaríamos de convidá-lo/a para analisar a ferramenta e emitir seu parecer técnico. Para isso, sugerimos que siga o seguinte roteiro:

1. Acesso: <https://prequest.websiteseguro.com/bi>
2. Avalie a relevância da solução proposta e se ela atende os objetos propostos ou não;
3. Avalie a ferramenta e seu potencial de adoção por um gestor;
4. Sua viabilidade comercial;
5. Sugira/critique a ferramenta com base nas métricas disponíveis lembrando que são sugestões iniciais e que podem ser adequadas às necessidades dos gestores;
6. Aponte sugestões de melhoria e inovação para a ferramenta;
7. Faça uma comparação desta ferramenta com outras soluções semelhantes existentes.

Agradecemos desde já sua atenção e disposição em contribuir com nossa pesquisa. Ficamos à disposição para qualquer esclarecimento adicional e aguardamos seu retorno até o dia xx de xxxx de 20xx.

Atenciosamente,  
Cleber Nauber dos Santos  
Instituto Federal de Alagoas – IFAL  
Doutorando PEI/UFBA