



UNIVERSIDADE FEDERAL DA BAHIA  
FACULDADE DE COMUNICAÇÃO  
PROGRAMA DE PÓS-GRADUAÇÃO EM  
COMUNICAÇÃO E CULTURA CONTEMPORÂNEAS

**FACOM**  
FACULDADE DE COMUNICAÇÃO - UFBA

**JORNALISMO DIGITAL EM BASE DE DADOS (JDBD)**  
**Um paradigma para produtos jornalísticos digitais dinâmicos**

**SUZANA BARBOSA**

SALVADOR - BAHIA

2007

S U Z A N A   B A R B O S A

**JORNALISMO DIGITAL EM BASE DE DADOS (JDBD)**  
**Um paradigma para produtos jornalísticos digitais dinâmicos**

Tese apresentada ao Curso de Doutorado em Comunicação e Cultura Contemporâneas, da Faculdade de Comunicação, da Universidade Federal da Bahia, como requisito parcial para a obtenção do grau de Doutora.

Orientador: Prof. Dr. Marcos Silva Palacios

SALVADOR - BAHIA

2007

**Biblioteca Central Reitor Macêdo Costa - UFBA**

B238 Barbosa, Suzana.

Jornalismo digital em base de dados (JDBD): um paradigma para produtos jornalísticos digitais dinâmicos / Suzana Barbosa. - 2007.

329 f. : il. + anexos.

Orientador : Prof. Dr. Marcos Silva Palacios.

Tese (doutorado) - Universidade Federal da Bahia, Faculdade de Comunicação, 2007.

1. Jornalismo. 2. Mídia digital. 3. Jornais eletrônicos. 4. Sites da web 5. Banco de dados. 6. Mineração de dados (Computação). 7. Metadados. 8. Metáfora. I. Palacios, Marcos Silva. II. Universidade Federal da Bahia. Faculdade de Comunicação. III. Título.

CDD - 070.4  
CDU - 070



**UNIVERSIDADE FEDERAL DA BAHIA**  
**FACULDADE DE COMUNICAÇÃO**  
**PROGRAMA DE PÓS-GRADUAÇÃO EM**  
**COMUNICAÇÃO E CULTURAS CONTEMPORÂNEAS**

**ATOS DE EXAME COMPREENSIVO DE DEFESA DE TESE**

**DOCTORANDA: SUZANA OLIVEIRA BARBOSA**

**TÍTULO DA TESE: “JORNALISMO DIGITAL EM BASE DE DADOS (JDBD) - Um paradigma para produtos jornalísticos digitais dinâmicos”**

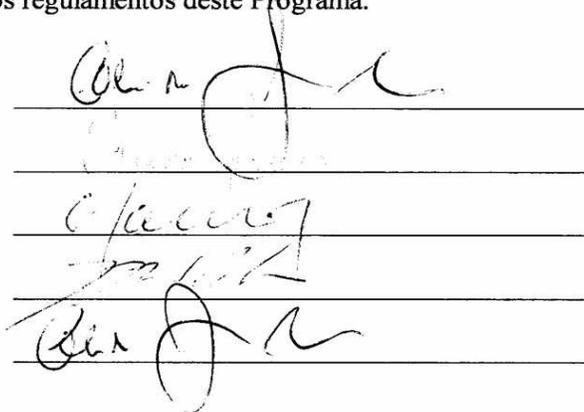
**DATA DO EXAME: 06 de março de 2007**

**EXAMINADORES:**

Prof. Dr. Cláudio Guimarães Cardoso (FACOM/UFBA);  
Profª Drª. Cláudia Irene de Quadros (Universidade Tuiuti/PR);  
Prof. Dr. Elias Machado Gonçalves (UFSC);  
Prof. Dr. Othon Fernando Jambeiro Barbosa (ICI/UFBA) e  
Prof. Dr. Marcos Silva Palacios (Orientador).

**PARECER COMPREENSIVO:**

Depois de avaliarmos a aula pública da Tese intitulada: “**JORNALISMO DIGITAL EM BASE DE DADOS (JDBD) - Um paradigma para produtos jornalísticos digitais dinâmicos**”, depositada no Curso de Doutorado deste Programa de Pós-Graduação, e a nós submetida para exame, e depois de realizados os ritos acadêmicos da defesa de tese, em que a doutoranda apresentou sua pesquisa e respondeu as nossas críticas, nós, os examinadores, decidimos, em sessão privada, que a doutoranda deve ser considerada APROVADA no Exame Compreensivo de Tese a que se submeteu em conformidade com os regulamentos deste Programa.



Salvador, 06 de março de 2007.

Aos meus pais  
– Eluísa Oliveira Barbosa e Zuarde Neiva Barbosa –  
e à minha família.

## AGRADECIMENTOS

*“(...) Eu aceito o belo presente  
Quero dar-lhe forma em silêncio  
Fazer desabrochar todas as suas cores  
E - sorrindo e desajeitado - ofertá-lo a ti (...)”*

- Rainer Maria Rilke.

Um tanto silenciosa – e serendipiosamente – esta tese foi aos poucos tomando forma. No caminho que trilhamos, alguns momentos foram mesmo penosos e outros tantos elucidadores, motivadores para persistirmos na descoberta e dar a conhecer as proposições que agora apresentamos. Em nossa trajetória, pudemos contar com o apoio, com o respeito e com o afeto de muitas pessoas. A elas, sou profundamente grata.

Ao poder superior, presente nas situações mais difíceis, concedendo-me luz e serenidade para superar inseguranças e prosseguir com coerência.

Ao meu pai, à minha mãe, à minha querida família, pela solidez, desvelo, compreensão e por ter criado as condições para que eu chegasse até aqui.

Ao Prof. Dr. *Marcos Palacios*, pela competência, confiança, encorajamento, paciência e gentileza. Um mestre admirável, com o qual aprendemos lições para a vida.

Ao Prof. Dr. *Elias Machado*, pelo estímulo, pela consideração e pelas contribuições relevantes.

Ao CNPq e à CAPES, pela oportunidade de dedicação exclusiva ao doutorado.

À coordenação, aos professores e aos funcionários do Programa de Pós-Graduação em Comunicação e Cultura Contemporâneas. Em especial: André, Carmem e Jeder.

Aos colegas do Grupo de Pesquisa em Jornalismo On-line (GJOL), pelo ambiente de investigação privilegiado, sempre marcado por idéias e discussões enriquecedoras.

À *Claudete Alves*, do Instituto de Matemática e CPD da UFBA, pelo valioso auxílio no início do desenvolvimento deste trabalho.

Ao Prof. Dr. *António Fidalgo*, pelo incentivo, apreço e atenção durante a realização do estágio doutoral sanduíche no Laboratório de Comunicação On-line (Labcom) da Universidade da Beira Interior, Covilhã - Portugal. Aos Profs. Drs. *Paulo Serra* e *Anabela Gradim*, e aos funcionários do Departamento de Artes e Letras da UBI. A *João Canavilhas*, pela acolhida e pela boa interlocução estabelecida.

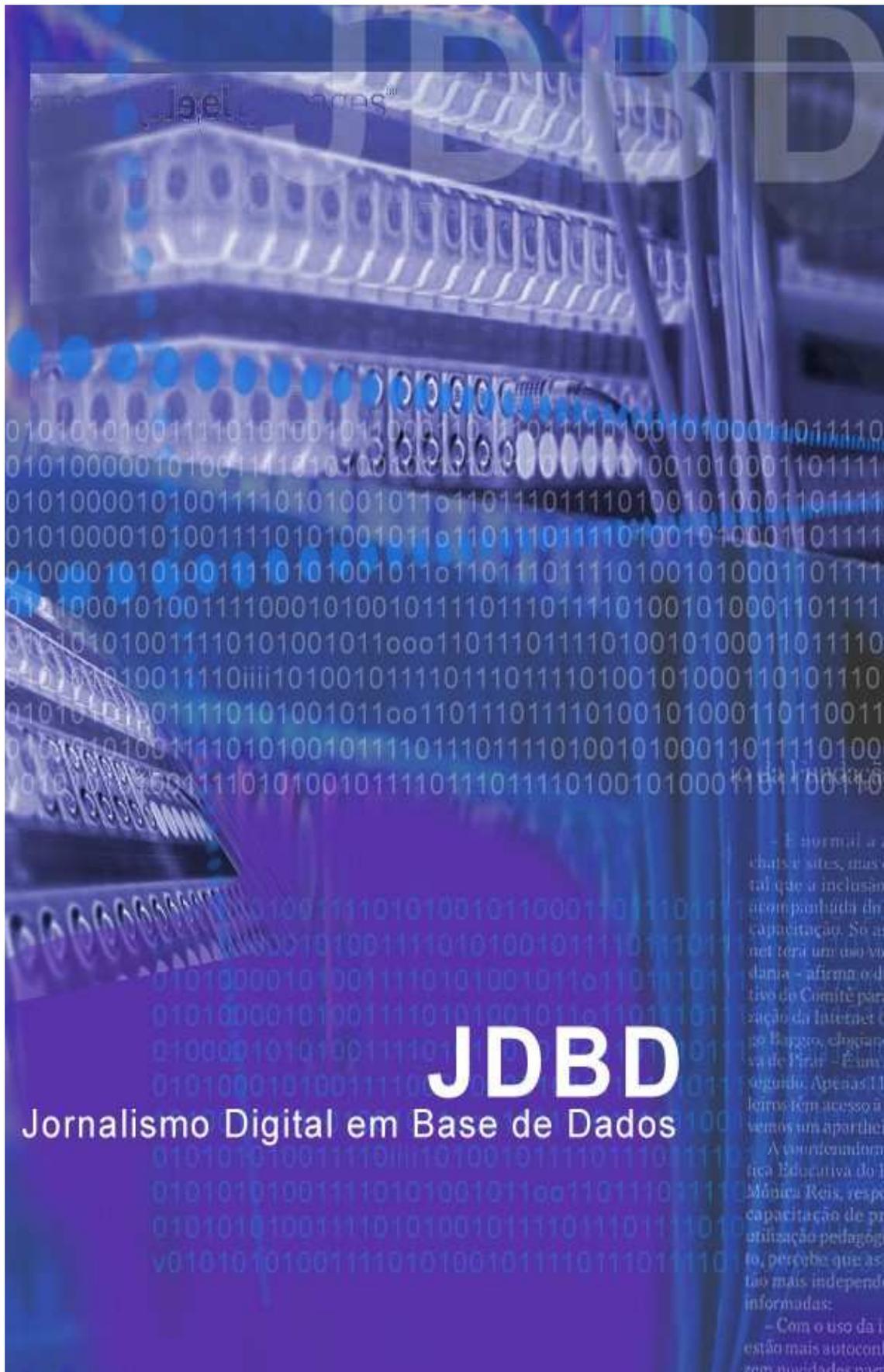
À *Jim Hall* (University College Falmouth, Reino Unido), *Concha Edo* (Universidad Complutense de Madrid), e *Javier Díaz Noci* (Universidad del País Vasco), pelo profícuo intercâmbio.

Aos amigos e colegas portugueses: *Catarina Rodrigues*, *Ivone Ferreira*, *Marta Filipa*, *Nélia Sousa*, *Catarina Moura*, *Eduardo Alves*, *Marco Oliveira* e *João Sardinha*, pela cordialidade e benquerença. À *Ana Jaleko* e *Augusto*, pela hospitalidade e alegria no melhor estilo luso-brasileiro. “Lisboa já tem sol, mas cheira a lua ...”

Aos amigos de sempre, e àqueles cuja amizade foi recentemente conquistada. Obrigada pelo carinho, pela torcida, solidariedade e pela disposição em ouvir: *Luciana Mielniczuk*, *Beatriz Ribas*, *Carla Schwingel*, *Lia Seixas*, *Ana Paula Prado*, *Eliane Pinho*, *Eliane Cardoso*, *Alexandra Dumas*, *Adriana Dumas*, *Jan Alyne*, *Mari Fiorelli*, *Afonso Júnior*, *Gil Maciel*, *Marlene Lopes*, *Cláudio Xavier*, *Joceval Santana*, *Danilo*, *Beto*, *Fabinho*, *Chico Vivas*, *Eledir*, *William*, *Zé Carlos*, *Bartira Bastos*, *Graciela Natanshon*, *José Mamede* e *Regina*, *Suzy dos Santos*, *Anna Jatobá*, *Lula Cardoso*, *Adriana Braga*, *Marcus Minuzzi*, *Raycho Mukelov*.

Ao meu irmão *Zuarte*, pelo cuidado, amabilidade e generosidade. À *Estelita* e à *Dona Antônia*, pela cooperação e bom humor.

À *Rumen* - moia lubov. Pela beleza do encontro. Por todo amor, força e inspiração. Obicham te!



# JDBD

## Jornalismo Digital em Base de Dados

- É normal a a  
chats e sites, mas  
tal que a inclusão  
acompanhada de  
capacitação. Só as  
nel terá um uso vi  
slaria - afirmou o di  
tivo do Comitê para  
zação da Internet (C  
go Baggio, coordena  
va de Pirat - É um t  
segundo. Apenas 11  
leiros têm acesso à  
vemos um apartidei

- A comunicação  
tica Educativa do P  
Mônica Reis, respo  
capacitação de pr  
ofissionalização pedagó  
to, percebe que as  
ção mais independe  
informadas.

- Com o uso da in  
estão mais autoconf  
zem novidades para

## RESUMO

Esta tese tem como preocupação central identificar e compreender os usos e as transformações que a crescente utilização de bases de dados está fazendo emergir para o jornalismo digital, modalidade em desenvolvimento há mais de uma década. No argumento endereçado, as bases de dados são entendidas como um aspecto-chave, inspirando a designação de um paradigma numa etapa de transição entre a terceira geração e uma quarta fase de evolução para o jornalismo digital, a saber: o **Jornalismo Digital em Base de Dados (JDBD)**, conformado segundo funcionalidades e categorias específicas. Reconhecemos, em paralelo, o seu potencial para fazer emergir uma nova metáfora para a apresentação dos conteúdos jornalísticos. A proposição de tal modelo é justificada, sobretudo, em função da complexificação dos processos para a implementação de produtos jornalísticos digitais que vem sendo processada nos últimos anos. Para a cartografia do objeto, elegemos um *corpus* empírico constituído por 22 casos, que ilustram e fundamentam as argumentações. Aliado ao método do estudo de caso, trabalhamos também dentro da perspectiva do método histórico, já que buscamos analisar o emprego de bases de dados no jornalismo através do tempo, para compreender a evolução do uso desta tecnologia até o momento atual. Assim, a tese explora o momento atual, indicando rupturas, identificando remediações, enxergando potencialidades, ao mesmo tempo em que se conecta com o processo histórico de emprego das bases de dados no jornalismo. O passado ajuda a melhor compreender o presente, a sua complexidade, e permite apontar para o futuro.

### **PALAVRAS-CHAVE:**

Jornalismo digital – Internet - Base de Dados – Paradigma - Metáfora.

## ABSTRACT

The main concern of this thesis is to identify and understand the uses and transformations the increasing exploitation of database has brought to surface for digital journalism, which is under construction for more than a decade. In the argument addressed by this thesis, databases are understood as a key feature, and inspire the designation of a paradigm in a transition stage between the third generation and the fourth phase of evolution for the digital journalism, namely: **Digital Journalism in Databases (DJDB)**, configured according to specific functionalities and categories. At the same time, we recognize its potential to bring a new metaphor for the presentation of journalistic content. The proposition of such model is justified, above all, by reason of the increasing complexity in the implementation processes of digital journalism products which have been processed in the last years. In order to identify the subject matter we selected an empirical *corpus* composed of 22 cases which illustrate and validate the argumentation. Besides applying the method of case study, we have worked in the perspective of the historical method as well, since we were seeking to analyze the use of databases in journalism along the years to understand the evolution of this technology usage up to present time. In this sense, the thesis exploits the present moment, indicating breakages, identifying kinds of remediation, pointing out potentialities while connecting, at the same time, with the historical process about the application of databases in journalism. The past helps, for a better understanding, the present and its complexity, and allows us to point to the future.

### KEY WORDS:

Digital Journalism – Internet – Database – Paradigm - Metaphor.

## LISTA DE FIGURAS

<b>Figura 1</b> – <i>Projeto File Room</i> , de Antonio Muntades, 1994-2004: casos de censura ao redor do mundo. (Out.2006)	72
<b>Figura 2</b> – <i>Time capsule</i> , evento-instalação de Eduardo Kac: corpo como interface base de dados. (Out.2006)	73
<b>Figura 3</b> – Instalação e projeto web, <i>Datamining Bodies in Ruhr</i> , idealizado por Victoria Vesna. (Out.2006)	74
<b>Figura 4</b> – <i>The Secret Life of Numbers</i> : popularidade dos números associados a vários eventos, desde 1997. (Out.2006)	75
<b>Figura 5</b> – <i>Soft Cinema</i> . Uma das cenas de <i>Mission to Earth</i> : alegoria sobre a imigração na Guerra Fria. (Out.2006)	77
<b>Figura 6</b> – <i>The Dumpster</i> : estratégias de análise de dados mostram novos tipos de representação social. (Out.2006)	78
<b>Figura 7</b> – <i>Newsmap</i> : demonstração visual de notícias geradas pelo <i>Google News</i> . (Out.2006)	79
<b>Figura 8</b> – <i>Newzbubble</i> . Visualização de informações jornalísticas através de interface de ‘bolhinhas’. (Out.2006)	80
<b>Figura 9</b> – News Brief do European Media Monitor. (Out.2006)	80
<b>Figura 10</b> – No <i>News Explorer</i> , por exemplo, visualiza-se a rede de relações em torno do presidente Lula. (Out.2006)	81
<b>Figura 11</b> - No mapa do <i>Deu no Jornal</i> , se deve entrar com os nomes das pessoas ou assuntos sobre os quais se quer visualizar as relações. (Dez.2006)	82
<b>Figura 12</b> – Seção <i>Most Popular</i> da <i>BBC News</i> usa técnicas diferenciadas para apresentar informações. (Out.2006)	82
<b>Figura 13</b> – Relação de semelhança entre os elementos	92
<b>Figura 14</b> – Um dos terminais exibindo informações disponibilizadas pela empresa de BDs, <i>LexisNexis</i> . Fonte: <a href="http://www.lexisnexis.com">www.lexisnexis.com</a> . (Set.2006)	105
<b>Figura 15</b> – O <i>Minitel 1</i> foi um dos aparelhos utilizados para se acessar o sistema desenvolvido na França. Fonte: <a href="http://en.wikipedia.org/wiki/Videotex">http://en.wikipedia.org/wiki/Videotex</a> . (Out.2006)	108

- Figura 16** – Tela da Central de Videotexto do Sindicato dos Jornalistas de Santos/SP, a primeira central de videotexto fora da capital paulista.  
Fonte: <http://www.novomilenio.inf.br/ano97/97hista7.htm>. (Out.2006) 109
- Figura 17** – *Home* do *Chicago Tribune*, disponibilizado via *America Online*, em 1992.  
Fonte: [http://iml.jou.ufl.edu/carlson/history/ScreenShots/chicago\\_online.jpg](http://iml.jou.ufl.edu/carlson/history/ScreenShots/chicago_online.jpg). 111
- Figura 18** – Em 1996, o *site* do *JB* já oferecia notícias em “tempo real”, com a criação da editoria “Extra”.  
Fonte: <http://jbonline.terra.com.br/destaques/2005/10anos/evolucao.html>. 112
- Figura 19** – Modelo da curva de adoção de inovação ou Curva do “S” explica o processo de difusão. Fonte: Rogers (1995) 115
- Figura 20** – Tela da *home* de artigos da *TDC*. Fonte: Imagem cedida por Raymond Colle (Out.2005) 159
- Figura 21** - Revista *Notibits de Informática Y Comunicaciones Digitales* ainda funciona sem sistema de metainformação (Out.2006) 160
- Figura 22** - O *Akademia* foi o primeiro produto do *Labcom* estruturado em base de dados (Maio.2006) 164
- Figura 23** – Reformulação permitiu ao *Urbi* deixar de ser estático para se tornar um produto dinâmico. (Jun.2006) 167
- Figura 24** – *Faces of the Fallen* do *Washingtonpost.com*: automatização para a apresentação do conteúdo. (Nov.2006) 176
- Figura 25** – *Votes Database*, também do *Washingtonpost.com*, agrega todas as informações sobre votações ocorridas no congresso dos EUA. (Nov.2006) 177
- Figura 26** – *Google News* é um dos mais conhecidos *sites* agregadores de notícias, funcionando de modo completamente automatizado. (Maio.2006) 179
- Figura 27** – O *Topix.net*, criado em 2002, tem tecnologia própria para oferecer conteúdos locais, o foco da sua oferta informativa. (Dez.2006) 180
- Figura 28** – Através de técnica de visualização específica, o *Newsmap* permite ver e comparar os assuntos que são destaque na cobertura de diversos *sites* jornalísticos, de diferentes países. (Jan.2007) 181
- Figura 29** – *Folha Online* é um dos poucos *sites* nacionais que exibem na *home* os “+Lidos”. (Dez.2006) 182
- Figura 30** – *Portal Terra* – Notícias: o subcanal “Fatos Históricos”. (Nov.2006) 183
- Figura 31** – O conteúdo do *blog Deu no Jornal* também pode ser lido no celular. (Dez.2006) 184
- Figura 32** – *BBC News* agrega as “memórias” de acontecimentos enviados por usuários, além de disponibilizar os relatos dos eventos a partir dos seus correspondentes espalhados pelo mundo. (Nov.2006) 185

- Figura 33** – No *Libération*, o arquivo está integrado às matérias, oferecendo contextualização e aprofundamento. Também se trabalha com o inter-relacionamento. (Jan.2007) 185
- Figura 34** – “Faces of the Dead” do *The New York Times*: interface “imagem mosaico” para contar histórias dos soldados mortos no Iraque. (Jan.2007) 187
- Figura 35** – O *The New York Times* também incorporou opções para agregar seu conteúdo a “sites sociais”. (Dez.2006) 189
- Figura 36** – A seta indica recursos para alterar a configuração do texto, traduzi-lo para idiomas distintos, compartilhá-los com sites da web 2.0, e permitem, também, escolher o número de colunas que se deseja ler o *International Herald Tribune*. As opções dinamizam o produto, que, por outro lado, mantém a referência da metáfora do impresso, com a leitura das matérias em duas ou três colunas. (Nov.2006) 189
- Figura 37** – A caixa com as opções de ler mais notícias relacionadas a um determinado assunto seja no *Elmundo.es* ou em outros sites é colocada ao final das matérias e no canto direito. Ao clicar, o usuário é levado à página com os resultados indexados. (Set.2006) 190
- Figura 38** – Página mostra lista de links para se ler mais sobre, por exemplo, assuntos relacionados ao papa Bento XVI. *Elmundo.es*. (Set.2006) 191
- Figura 39** – *Elpais.com*. Seta indica caixa com opções de inter-relacionamento. Abaixo, os ícones para compartilhamento. (Dez.2006) 191
- Figura 40** – Modelo proposto para criar produtos com formatos distintos, a partir de templates, no novo *Panopticon*. (Dez.2006). 198
- Figura 41** – Página do publicador *MapaLink* com a representação da notícia hipertextual. Fonte: MIELNICZUK; MARQUES, 2006, p.14. 201
- Figura 42** – Categorias do Jornalismo Digital em Base de Dados (JDBD) 233
- Figura 43** – *DailySource*: agregador incorpora jornalistas para o trabalho de seleção e edição e não possui anúncios publicitários. (Jan.2007) 253
- Figura 44** – O *Daylife* promete maneira diferente para explorar o mundo, com perspectivas variadas e contextualização. (Fev.2007) 254
- Figura 45** – Técnica da pirâmide deitada contempla as BDs nos níveis de contextualização e exploração, linkando para os arquivos da publicação e para sites externos. Fonte: Canavilhas (2005, 2006). 256
- Figura 46** – Gestor de conteúdos do *Elmundo.es*. Marcadores indicam os botões de inserção para links de inter-relacionamento apontando para o arquivo do produto, bem como para o de outros sites. Fonte: Rivero, 2006. 261

**LISTA DE QUADROS**

<b>Quadro 1</b> – As fases da pesquisa	40
<b>Quadro 2</b> - Princípios, estágios e níveis de participação do usuário	95
<b>Quadro 3</b> – Características das Inovações	119
<b>Quadro 4</b> – Gerações de desenvolvimento do Jornalismo Digital	148
<b>Quadro 5</b> – JDBD: novo paradigma surge entre a 3 <sup>a</sup> e a 4 <sup>a</sup> geração	150
<b>Quadro 6</b> – JDBD: Conceito, Funcionalidades e Categorias	220
<b>Quadro 7</b> – Elementos conceituais que organizam a classificação externa dos conteúdos, tecnologias e recursos agregados à oferta informativa	266

## LISTA DE ABREVIATURAS E SIGLAS

Ajax – Asynchronous JavaScript and XML

ASP - Active Server Page

BDs - Base de Dados

CAJ - Computer Assisted Journalism

CAR - Computer Assisted Reporting

CMS - Content Management Systems

CSS – Cascading Style Sheets

EMM – European Media Monitor

GJOL – Grupo de Pesquisa em Jornalismo On-line

GUI - Graphical User Interface

HCI - Human-Computer Interface

HTML - Hypertext Markup Language

IRE – Investigative Reporters and Editors

JAC - Jornalismo Assistido por Computador

JDBD - Jornalismo Digital em Base de Dados

KDD - Knowledge Discovery in Databases

Labcom – Laboratório de Comunicação On-line

MP3 - MPEG Audio Layer-3

MySQL – Software Livre de Gerenciamento de Base de Dados, baseado em SQL

NICAR – National Institute for Computer-Assisted Reporting

PC - Personal Computer

PDA – Personal Digital Assistant

PDF - Portable Document Format

PHP – Hypertext Preprocessor

PIP – Produto de Implementação do Panopticon

PPEP - Plataforma de Publicação e Ensino Panopticon

RAC - Reportagem Assistida por Computador

RSS – Really Simple Syndication ou Rich Site Summary

SGBD - Sistema de Gerenciamento de Base de Dados

SQL – Structured Query Language

TICs – Tecnologias da Informação e Comunicação

TDC – Temas Digitales de Comunicación

UBI – Universidade da Beira Interior

VDTs - Video Display Terminals

XML – eXtensible Markup Language

WWW - World Wide Web

W3C – World Wide Web Consortium

3D - Três Dimensões

## SUMÁRIO

<b>DEDICATÓRIA</b>	v	
<b>AGRADECIMENTOS</b>	vi	
<b>EPÍGRAFE</b>	viii	
<b>RESUMO</b>	ix	
<b>ABSTRACT</b>	x	
<b>LISTA DE FIGURAS</b>	xi	
<b>LISTA DE QUADROS</b>	xiv	
<b>LISTA DE ABREVIATURAS E SIGLAS</b>	xv	
<b>INTRODUÇÃO</b>	20	
I	Por que estudar o uso de base de dados no jornalismo digital	21
II	Definindo o Jornalismo Digital em Base de Dados – “JDBD”	26
III	Objetivos e hipóteses	29
IV	Referencial teórico	31
V	Delimitação do objeto e casos estudados	36
VI	Metodologia	39
VII	Estrutura da tese	41
<b>1</b>	<b>A FORMA CULTURAL BASE DE DADOS</b>	44
<b>1.1</b>	<b>O “Lugar” das Bases de Dados na Contemporaneidade</b>	45
1.1.1	Como uma tecnologia passa a “forma cultural”	53
1.1.2	Web como base de dados	59
<b>1.2</b>	<b>Interface para Produtos da Nova Mídia</b>	63
1.2.1	Uma estética: modos diferenciados para visualizar a informação	66
1.2.2	Base de dados como ferramenta de comunicação	83
1.2.3	Uma hipernarrativa: a narrativa interativa	86
<b>2</b>	<b>AS BASES DE DADOS E O JORNALISMO</b>	97
<b>2.1</b>	<b>Histórico de Adoção Segue Curva do “S”</b>	98
2.1.1	Abordagem ancorada na teoria da difusão de inovações	114

<b>2.2</b>	<b>Emprego de Bases de Dados nas Empresas Informativas</b>	123
2.2.1	Estrutura para armazenamento de informações	125
2.2.2	Sistemas para entrega eletrônica de notícias	128
2.2.3	Fontes para coleta, apuração e recuperação de informações	131
2.2.4	A segunda geração da <i>CAR</i> e a “ <i>database era</i> ”	134
2.2.5	As bases de dados e as primeiras versões de <i>web sites</i> jornalísticos	139
<b>3</b>	<b>O QUE É JORNALISMO DIGITAL EM BASE DE DADOS – (JDBD)?</b>	144
<b>3.1</b>	<b>Um Paradigma na Transição da 3ª para a 4ª Geração</b>	145
3.1.1	Atribuição de um novo <i>status</i> : a remediação e a extensão do conceito	153
3.1.2	Aporte teórico e o estado da arte das bases de dados no jornalismo	157
3.1.2.1	O valor da informação estruturada e a automatização dos processos	173
3.1.3	Sistemas de gestão de conteúdos e as bases de dados	193
3.1.4	Tipos de bases de dados	203
<b>4</b>	<b>DELIMITANDO O “JDBD” E AS SUAS PROPRIEDADES</b>	209
<b>4.1</b>	<b>Dimensão Conceitual e os Elementos de Especificidade</b>	210
4.1.1	A definição e a concepção operacional de “JDBD”	214
<b>4.2</b>	<b>As Funcionalidades</b>	221
4.2.1	Descrevendo as 18 funções identificadas	222
<b>4.3</b>	<b>As Categorias</b>	232
4.3.1	Dinamicidade	233
4.3.2	Automatização	236
4.3.3	Flexibilidade	237
4.3.4	Inter-relacionamento/Hiperlinkagem	238
4.3.5	Densidade informativa	239
4.3.6	Diversidade temática	243
4.3.7	Visualização	244
<b>5</b>	<b>UMA NOVA METÁFORA A TOMAR FORMA</b>	245
<b>5.1</b>	<b>Rupturas Qualitativas e Potencialidades do “JDBD”</b>	246
5.1.1	Produto digital metainformativo, inteligente e dinâmico	249
5.1.2	Sistemas de produção, edição e circulação automatizados	251

5.1.3	Composição das peças informativas e as notícias de ordem superior	254
5.1.4	Arquivos como “bases de conhecimento”	259
5.1.5	Evolução nos modos de apresentação das informações	262
5.1.6	Elementos conceituais para a organização do conteúdo	264
5.1.7	O jornalista, as competências cognitivas, as novas funções	267
5.1.8	Usuários atuando como fonte informativa	271
<b>CONCLUSÕES</b>		276
<b>REFERÊNCIAS BIBLIOGRÁFICAS</b>		285
<b>GLOSSÁRIO</b>		317
<b>APÊNDICES</b>		327
<b>APÊNDICE I – Produtos que compõem o <i>corpus</i> empírico</b>		328
<b>APÊNDICE II – Formulário de Observação</b>		329

## **INTRODUÇÃO**

## Introdução

---

### I - Por que estudar o uso de base de dados no jornalismo digital

Da condição quase apenas perceptível para os informatas durante muitas décadas, as bases de dados (BDs), no ambiente complexo da sociedade contemporânea, estendem a sua visibilidade a diversas áreas, por conta dos usos e práticas que ajudam a forjar. Já é notável que, em algumas delas, o emprego de tal tecnologia da informação ocorre segundo novas perspectivas e abordagens alinhadas à concepção de base de dados como uma forma cultural simbólica. Uma forma característica de uma era na qual o computador é o amplificador social das descobertas e o elemento constitutivo de práticas.

Para o jornalismo digital – modalidade jornalística em desenvolvimento há mais de uma década –, as bases de dados vão inspirar a designação de um paradigma, pelas especificidades e potencialidades que geram, a saber: o **Jornalismo Digital em Base de Dados (JDBD)**.

A denominação **Jornalismo Digital em Base de Dados** desponta em razão das funcionalidades asseguradas pelas bases de dados para a construção e gestão de produtos jornalísticos digitais, bem como para a estruturação e a apresentação dos conteúdos. A realização desta tese nos coloca, assim, ante o desafio de olhar para algo cujas implicações concretas para esta modalidade de jornalismo começam a se fazer conhecidas, compreendidas.

Chegamos a essa designação com o progressivo desenvolvimento da pesquisa, mais exatamente em setembro de 2005, quando cumpríamos o período do estágio doutoral sanduíche, no Laboratório de Comunicação On-line (Labcom) da Universidade da Beira

Interior (UBI), Portugal. O título foi parcialmente inspirado no *paper* “*What is Multimedia Journalism*”?, de Mark Deuze (2004).

A primeira vez que expusemos a idéia do Jornalismo Digital em Base de Dados foi nas apresentações realizadas nas Mesas “Jornalismo” e “Novas Tecnologias, Novas Linguagens”, durante o IV Congresso da Associação Portuguesa de Ciências da Comunicação – Sopcom –, realizado na Universidade de Aveiro (Aveiro, PT), nos dias 20 e 21 de outubro de 2005. Naqueles textos (BARBOSA, 2005a, 2005b), tratávamos de um conceito operacional para as bases de dados no jornalismo digital, além de algumas funcionalidades e noções características. Em eventos que se seguiram, como as Jornadas *Jornalismo Online.2005: Aspectos e Tendências* (UBI, Covilhã, PT, 2005), *VII Lusocom* (Universidad de Santiago de Compostela, Espanha, 2006) e *XV Encontro Anual da Compós* (Bauru, SP, 2006), participamos com trabalhos específicos sobre este assunto.

No argumento endereçado por esta investigação, perpassa a idéia de cisão, de ruptura no significado das bases de dados, no seu *status* e no modo como até então se apreendia o seu emprego. Atualmente, muitos estudos – circunscritos tanto na área das novas mídias, da cibercultura, da comunicação, do jornalismo, das artes, das ciências da computação, da inteligência artificial, da economia – têm delineado perspectivas e apontado funções para as bases de dados que descortinam as novas linhas de força, os modos de operação, de representação, as potencialidades que as atravessam e as colocam num patamar diferenciado.

Por isso mesmo, a escolha da presente temática foi guiada pelo propósito de contribuir para o melhor entendimento acerca das transformações que a crescente utilização de bases de dados está fazendo emergir para o jornalismo digital. Ao circunscrever o estudo numa etapa de transição entre a terceira geração (MIELNICZUK, 2003; PRYOR, 2002) e uma quarta fase de evolução para esta modalidade jornalística, destacamos que o recorte se faz necessário porque estamos considerando as bases de dados como um paradigma neste momento, além de

estarmos trabalhando com a perspectiva de produtos noticiosos digitais já mais descolados daqueles que marcaram as etapas anteriores, distinguidas pela transposição de conteúdos, uso limitado de recursos propiciados pelo suporte digital e conhecimento parcial das características (hipertextualidade, multimídia, interatividade, personalização, memória/arquivo) que conformam o jornalismo digital como uma modalidade específica para o jornalismo – depois da imprensa, do cinema, do rádio e da televisão.

Esta fase de transição é marcada por uma base tecnológica ampliada, acesso expandido por meio de conexões banda larga; proliferação de plataformas móveis; equipes mais especializadas; uso expandido de bases de dados; algoritmos; linguagens de programação; desenvolvimento de sistemas de gestão de conteúdos mais complexos; maior incorporação dos *blogs*; adoção de sistemas que habilitem a participação efetiva do usuário na produção de informações; produtos diferenciados criados e mantidos de modo automatizado; *sites* dinâmicos; narrativas multimídia; utilização de recursos como *RSS (Really Simple Syndication* ou *Rich Site Summary*) para recolher, difundir e compartilhar conteúdos; uso da técnica do *podcasting* para distribuição de conteúdos em áudio e em vídeo; experimentação de novos elementos conceituais para a organização da informação; maior integração do material de arquivo na oferta informativa; emprego de metadados e *data mining* para extração de conhecimento; e aplicação de novos métodos para gerar visualizações diferenciadas para os conteúdos jornalísticos.

Neste cenário, o modelo que propomos é o **Jornalismo Digital em Base de Dados (JDBD)**, conformado segundo funcionalidades e categorias específicas. Ao apontarmos tal modelo como paradigma de uma fase de transição, ressaltamos que o quarto estágio se encontra ainda em gestação. Porém, seja qual for a configuração que irá consolidá-lo, as bases de dados serão um elemento estruturante, pela flexibilidade intrínseca a elas, assim como o potencial de reinvenção que elas possuem.

Partimos do pressuposto que tudo isso se dá no contexto de uma sociedade informacional, computadorizada, marcada pelo liberalismo econômico (CASTELLS, 1999, 2001; HARVEY, 2000), e pela convergência tecnológica entre a informática, as telecomunicações e a microeletrônica, da qual resultaram as Tecnologias da Informação e Comunicação (TICs): internet, computadores, World Wide Web, telefones celulares, entre outros dispositivos). Nessa perspectiva, as TICs são vistas como tecnologias de integração pois constituem o futuro da convergência de várias tecnologias prévias e modificam o conjunto das mesmas, estando atualmente amplamente difundidas entre os diversos campos do saber (LEVINSON, 1998; LÉVY, 1993, 1999; ECHEVERRÍA, 1994, 1999; LIEVROUW; LIVINGSTONE, 2002; LIEVROUW, 2004; GRAHAM, 2004).

A motivação para estudar o uso de base de dados no jornalismo digital surgiu na fase inicial do doutoramento (maio/2003), ao cursar a disciplina optativa COM 543 (Ciberespaço, Comunicação e Cultura. Modelos de Narrativa e Gêneros no Jornalismo Digital), ministrada pelo professor Elias Machado Gonçalves. Foi a partir das aulas e do artigo produzido, assim como das discussões e atividades desenvolvidas no âmbito do Grupo de Pesquisa em Jornalismo On-line (GJOL), que germinou a idéia desta investigação.

Uma vez instigada a refletir sobre a “originalidade” de um objeto de pesquisa que se projetava, o desdobramento natural foi a decisão de proceder a uma total reformulação do projeto com o qual havia sido selecionada para o doutorado neste programa de pós-graduação. A mudança do projeto aconteceu, primordialmente, em razão do possível cariz inovador e da pertinência do estudo para o jornalismo digital.

De início, a abordagem esboçada evidenciava muito fortemente o caráter de ruptura e a primazia da tecnologia, felizmente contrabalançada, relativizada, na medida em que se avançava com a investigação. Embora, como afirmamos, a idéia de ruptura caracterize a argumentação da tese, por outro lado, coloca-se, em paralelo, a existência de remediações e

potencialidades, que, lado a lado com as mudanças ocorridas, a quebra de certos padrões e as novas concepções atribuídas, vão auxiliar na adequação da abordagem construída em torno do que chamamos **Jornalismo Digital em Base de Dados (JDBD)**.

Por outro lado, pesquisar tal assunto também nos aponta ou sugere outras razões e motivos pelos quais os pesquisadores têm estudado o uso de bases de dados no jornalismo desde a descoberta da aplicação potencial desta tecnologia para a área. Tudo começa efetivamente na década de 70, quando as bases de dados passam a ser empregadas, inicialmente, para o armazenamento e distribuição de informações jornalísticas; em seguida, são introduzidas no processo de apuração, para, numa etapa seguinte, se consolidarem como protagonistas no desenvolvimento da Reportagem Assistida por Computador ou *Computer Assisted Reporting (CAR)*.

O uso posterior vai estar vinculado ao videotexto, que desponta no início dos anos 80 – assim como o computador pessoal – e experimenta expansão no correr da década até o início dos anos 90 como a grande promessa de futuro para as organizações jornalísticas. No entanto, em meados da década de 90, o surgimento da World Wide Web (WWW) faz com que a direção das previsões mude de rota e aponte a internet – e o suporte digital – como o lugar de inovação, para onde tudo convergiria. Esse é o ponto de mudança a desvelar outro momento de expansão para o emprego das bases de dados, agora vinculadas a uma nova modalidade jornalística e que resultará naquilo que denominamos como **Jornalismo Digital em Base de Dados**.

Desta maneira, a tese explora o momento atual, indicando mudanças, identificando remediações, enxergando potencialidades, ao mesmo tempo em que se conecta com o processo histórico de emprego das bases de dados no jornalismo. O passado ajuda a melhor compreender o presente, a sua complexidade, e permite apontar para o futuro.

## II – Definindo o Jornalismo Digital em Base de Dados – “JDBD”

Inicialmente, ressaltamos que a definição de base de dados (BDs)<sup>1</sup> que colocamos em operação tenta contemplar uma percepção mais ampla sobre essa tecnologia da informação (DATE, 1984; PEREIRA, 1998; SOUSA, 2002; GUIMARÃES, 2003). Para isso, incorporamos como principais referências a concepção de Lev Manovich (2001), que as considera forma cultural simbólica na contemporaneidade, e o conceito trabalhado por Elias Machado (2004a, 2004b, 2006), que, baseado em Manovich, propôs a idéia de base de dados como formato no jornalismo digital.

Para a sua formulação, também levamos em conta os trabalhos de pesquisadores que têm estudado a aplicação desta tecnologia no jornalismo digital de modo mais direto ou indireto (COLLE, 2002, 2005a, 2005b; PALACIOS, 2002a, 2003b, 2003c, 2004; FIDALGO, 2003, 2004; LOPEZ, GAGO, PEREIRA, 2003; QUADROS, 2004, 2005; LIMA JÚNIOR, 2004, 2005, 2006; GARCÍA et al, 2005; HOLOVATY, 2005, 2006a, 2006b; GAGO, 2006), conectando-os aos de outros investigadores da área da comunicação e do jornalismo (KOCH, 1991; CARITÁ, 1994; PAUL, 1999; GARRISON, 1998; HERBERT, 2000; REAVY, 2001; PAVLIK, 2001, 2005; LAGE, 2001; 2002; QUINN, 2002, 2005; GUNTER, 2003; PRYOR, 2002; QUIJADA, 2006).

Além deles, agregamos estudos provenientes das ciências da computação (FRANGOMENI, 1987; SOL, 1998; MAURER, SCHERBAKOV, HALIM, RAZAK, 1998;

---

<sup>1</sup> Optamos pelo uso do termo “base de dados” neste trabalho para denominar tanto a estrutura lógica como o conteúdo. Notamos, a partir da revisão bibliográfica, que há um maior predomínio do termo “base de dados”, traduzido do inglês *database*. Há, contudo, autores que diferenciam “base”, como sendo a estrutura lógico-matemática, de “banco”, relacionado ao conteúdo, à informação, enquanto muitos outros autores empregam “base” para denominar ambas as esferas. De acordo com Raymond Colle (2002, p.30), o termo “banco de dados” tem caído em desuso, sendo substituído por *datawarehouse* (conjunto integrado de dados ou depósito de dados). Embora ele aponte a distinção entre “banco” e “base”, emprega BD como acrônimo para “base de dados”. O mais comum entre autores brasileiros é usar “banco de dados” ou ambos os termos como sinônimos, ao contrário dos autores portugueses e espanhóis, em cujos trabalhos há o predomínio de “base de dados”. Assim, apesar de optarmos por “base de dados”, em alguns momentos usamos “banco de dados”, como sinônimo.

SOUSA, 2002; PIATTINI; CALERO, 1999; PIATTINI, 2000; BRICKLIN, 2001; VILLEGAS Y GARCÍA, 2002; CARLEY, 2002; JORGENSEN, 2003; SALGADO, RODRÍGUEZ, 2003; PAUL, 2004; GRAY, 2004), e das ciências da informação (SEMONCHE, 1993a, 1993b; RECODER, ABADAL, CODINA, 1995; PUJOL, 1995; SELLARÉS, 1995; MARCOS RECIO, 1999; TRAMULLAS Y KRONOS, 1997, 2000).

Para o jornalismo digital, as bases de dados são definidoras da estrutura e da organização, bem como da apresentação dos conteúdos de natureza jornalística. Elas são o elemento fundamental na constituição de sistemas complexos para a criação, manutenção, atualização, disponibilização e circulação de produtos jornalísticos digitais dinâmicos.

Como tal, desempenham um conjunto de funções percebidas tanto quanto à gestão interna dos produtos, quanto aos processos de apuração e contextualização, à estruturação das informações, à composição das peças informativas, assim como à recuperação das informações e à apresentação dos conteúdos. São elas:

- Indexar e classificar as peças informativas e os objetos multimídia;
- Integrar os processos de apuração, composição e edição dos conteúdos;
- Conformar padrões novos para a construção das peças informativas;
- Agilizar a produção de conteúdos, em particular os de tipo multimídia;
- Propiciar categorias diferenciadas para a classificação externa dos conteúdos;
- Estocar o material produzido e preservar os arquivos (memória), assegurando o processo de recuperação das informações;
- Permitir usos e concepções diferenciadas para o material de arquivo;
- Garantir a flexibilidade combinatória e o relacionamento entre os conteúdos;
- Gerar resumos de notícias estruturados e/ou matérias de modo automatizado;
- Armazenar anotações semânticas sobre os conteúdos inseridos;

- Habilitar o uso de metadados para análise de informações e extração de conhecimento, seja por meio de técnicas estatísticas ou métodos de visualização e exploração, como o *data mining*;
- Ordenar e qualificar os colaboradores e “repórteres cidadãos”;
- Orientar e apoiar o processo de apuração, coleta e contextualização dos conteúdos;
- Regular o sistema de categorização de fontes jornalísticas;
- Sistematizar a identificação dos profissionais da redação;
- Cartografar o perfil dos usuários;
- Transmitir e gerar informação para dispositivos móveis (celulares, computadores de mão, *iPods*, entre outros);
- Implementar publicidade dirigida.

A delimitação das funções foi pensada a partir de algumas noções propostas, inicialmente, como características do **Jornalismo Digital em Base de Dados (JDBD)**, (2005a; 2005b; 2006a, 2006b), utilizadas posteriormente como modelos de apreensão do objeto em estudo, quais sejam: resolução semântica, metadados, relato imersivo ou narrativa multimídia e jornalismo participativo. Da associação entre as noções características – em especial a de resolução semântica – e as funcionalidades, elaboramos sete categorias descritivas, que alcançam objetividade dado o caráter operativo e tornam perceptíveis as especificidades do **JDBD**. Estas categorias são: a) Dinamicidade; b) Automatização; c) Inter-relacionamento/Hiperlinkagem; d) Flexibilidade; e) Densidade informativa; f) Diversidade temática; g) Visualização.

Assim, conceituamos **Jornalismo Digital em Base de Dados (JDBD)** como o modelo que tem as bases de dados como definidoras da estrutura e da organização, além da

apresentação dos conteúdos de natureza jornalística, de acordo com funcionalidades e categorias específicas, que vão permitir a criação, a manutenção, a atualização, a disponibilização e a circulação de produtos jornalísticos digitais dinâmicos. Além de um paradigma na transição entre a terceira e a quarta gerações do jornalismo digital, argumentamos que este modelo tem, por outro lado, o potencial para assegurar uma metáfora particular, relacionada com os modos diferenciados para a visualização das informações.

### **III – Objetivos e hipóteses**

Este trabalho buscou compreender o uso de bases de dados no jornalismo digital. Para tanto, definimos como objetivo central identificar as remediações, as rupturas e as potencialidades que a crescente utilização dessa tecnologia da informação gera para esta modalidade jornalística. Derivados do objetivo geral, estabelecemos alguns objetivos específicos. Foram eles:

1. Avaliar o histórico de adoção das bases de dados no jornalismo;
2. Verificar os aspectos relacionados às bases de dados como uma forma cultural do contemporâneo;
3. Indicar a emergência de um novo *status* para as bases de dados no jornalismo digital;
4. Construir um conceito operacional para as bases de dados no jornalismo digital;
5. Delimitar as propriedades (funcionalidades e categorias) das bases de dados no jornalismo digital;

6. Verificar a aplicação das funcionalidades e das categorias estabelecidas para as bases de dados nos produtos selecionados que integram o *corpus* empírico;
7. Identificar a existência de novos elementos conceituais para a apresentação das informações;
8. Perceber o potencial do uso de bases de dados para a criação de uma nova metáfora no jornalismo digital.

Juntamente com os objetivos, foi formulada a hipótese principal, a qual foi desdobrada em cinco hipóteses secundárias para que ficasse mais clara e assegurasse uma estruturação coerente para os capítulos. A sua elaboração surgiu nas reflexões e nos cruzamentos estabelecidos durante o processo de construção da abordagem teórica, bem como da observação do quadro empírico, o que implicou o refinamento de algumas delas no processo de desenvolvimento da tese. Temos, então, como hipóteses:

#### Principal

- As bases de dados produzem remediações, rupturas e potencialidades para o jornalismo digital no que se refere aos formatos dos produtos, isto é, a estruturação e a organização das informações, a construção das peças informativas, além da apresentação dos conteúdos;

#### Secundárias

1. O uso expandido das bases de dados faz emergir um novo *status* para essa tecnologia no campo do jornalismo;
2. As bases de dados são um aspecto-chave no atual estágio de desenvolvimento do jornalismo digital;

3. As bases de dados possibilitam a criação de produtos jornalísticos digitais dinâmicos;
4. As bases de dados conferem automatização e maior agilidade aos processos de produção, composição, edição, recuperação e apresentação das informações no jornalismo digital;
5. As bases de dados asseguram a criação de um novo paradigma para o jornalismo digital.

#### **IV. Referencial teórico**

Para identificar as rupturas e remediações que a crescente utilização de bases de dados gera para o jornalismo digital em seu atual estágio de desenvolvimento, foi necessário ir um pouco além dos limites do jornalismo e do jornalismo digital. Buscamos embasamento nos estudos sobre os meios, sobre a internet, a nova mídia, a teoria da difusão de inovações, a teoria literária, a teoria crítica, a web arte, os sistemas de informação, as tecnologias da informação e comunicação, as linguagens de programação, as técnicas de descoberta de conhecimento entre os dados e os métodos para a visualização de informações.

Desta maneira, construímos uma abordagem teórica fundamentada num marco conceitual abrangente, que nos permitiu um melhor enquadramento do objeto de modo a contemplar as hipóteses formuladas e delimitar o que denominamos **Jornalismo Digital em Base de Dados (JDBD)**. Um referencial teórico multidisciplinar nos permitiu, por outro lado, preencher algumas lacunas geradas pela ainda incipiente literatura específica sobre o objeto investigado.

Os primeiros estudos sistematizados sobre o emprego da tecnologia das bases de dados (BDs) no jornalismo são da década de 90. O entendimento sobre como se desenvolveram as

apropriações iniciais das bases de dados no jornalismo veio com as obras de Tom Koch (1991), *Journalism for The 21<sup>st</sup> Century. Online Information, Electronic Databases and the News*; Philip Meyer, *Periodismo de Precisión. Nuevas Fronteras Para La Investigación Periodística* (1993); Bruce Garrison (1998) *Computer-Assisted Reporting*; e Nora Paul, *Computer Assisted Research: A Guide to Tapping Online Information* (1999). Em geral, a literatura deste período coloca a tecnologia, o computador e as BDs como ferramentas, ou seja, desempenhando um papel auxiliar.

Também encontramos que, antes mesmo de Koch ter afirmado que a utilização crescente das bases de dados revolucionaria as redações no seu trabalho de coleta de informações, escrita e apresentação das notícias, coube a Anthony Smith, em *Goodbye Gutenberg. The Newspaper Revolution of The 1980s* (1980), perceber que a chegada dos computadores às redações favoreceria o uso de várias bases de dados *on-line* e mudaria o modo como os jornalistas coletavam e sintetizavam as notícias.

A utilização do computador como ferramenta e a WWW como a maior das BDs disponíveis, por meio das quais se passava a ter acesso a incontáveis bases de dados governamentais, comerciais, bem como as das próprias empresas jornalísticas para apuração, análise e aprofundamento, é o tema central de outros autores (BORDEN, HARVEY, 1998; HERBERT, 2001; REAVY, 2001; HOUSTON, 2004). Nesta mesma linha, somam-se outras importantes contribuições (CARITÁ, 1987; COLOMBO, 1991; VIANNA, 1992; SEMONCHE, 1993; SIQUEIRA, 1995; CODINA; RECODER, ABADAL, 1995; PUJOL, 1995; SELLARÉS, 1995; ARMAÑANZAS, NOCI, MESO, 1996; LANCASTER, 1997; MARCOS RECIO, 1999; BALDESSAR, 2003).

Está nos trabalhos recentes de investigadores que têm ampliado a dimensão conceitual para as bases de dados no jornalismo e, especificamente, no jornalismo digital (HALL, 2001; LAGE, 2001; 2002; PALACIOS, 2002a, 2003b, 2003c; 2004; PRYOR, 2002; LÓPEZ,

GAGO, PEREIRA, 2003; QUINN, 2002, 2005; FIDALGO, 2003, 2004, 2005; COLLE, 2002, 2005a, 2005b; MACHADO, 2004a; 2004b, 2006; LIMA JÚNIOR, 2004, 2005, 2006; QUADROS, 2004, 2005; PAVLIK, 2001, 2005; GARCÍA et al, 2005; HOLOVATY, 2005, 2006a, 2006b; GAGO, 2006) a fonte fundamental para a percepção de um novo *status* para as bases de dados. Também com esses autores, evidenciou-se a passagem de uma visão que percebia as novas tecnologias e o computador como ferramentas para outra que passa a considerá-los como elementos fundamentais e constitutivos de uma nova prática jornalística.

Junto aos conceitos, especificidades e funcionalidades apreendidos do estudo desses autores, incorporamos a concepção de bases de dados como forma cultural, proposta por Lev Manovich na obra *The Language of New Media* (2001) e que está presente em vários dos seus artigos, ensaios, *papers* e projetos artísticos (2000, 2003, 2004, 2005, 2006). A idéia foi aqui apropriada porque possibilita o entendimento acerca da ubiquidade da tecnologia das bases de dados na contemporaneidade e, especificamente, sobre o seu uso potencial no jornalismo digital.

Para Manovich, as bases de dados são um complexo de armazenagem de formas culturais e, como tal, estão no centro do processo criativo para a construção de produtos da nova mídia. Além disso, engendram um tipo específico de narrativa, a hipernarrativa, construída pela ligação de elementos da base de dados em uma ordem particular. Também é a partir de um dos princípios da nova mídia – transcodificação –, descritos por Lev Manovich, que Elias Machado (2004a, 2004b; 2006) vai lançar a hipótese das bases de dados como um formato no jornalismo digital. Tal idéia está, simultaneamente, sintonizada com a nossa proposição sobre o potencial que as bases de dados possuem para originar uma nova metáfora (BARBOSA, 2004a, 2004b, 2004c, 2005a, 2005b, 2006a, 2006b).

Para compreender melhor a concepção de BDs como uma forma cultural, recorreremos aos estudos de Raymond Williams (1977, 1990) e de Álvaro Vieira Pinto (2005), por um lado,

e, do outro, procuramos incorporar as análises mais recentes sobre o papel desta tecnologia e suas aplicações em diversos setores de atividade (SOL, 1998; VESNA, 1999, 2000; DIETZ, 2000, 2004; NOGUEIRA, 2001, 2003; LEÃO, 2003; PAUL, 2004; LUCAS, 2004, 2006).

A partir de Barrie Gunter (2003), pudemos identificar um novo momento ascendente para o emprego das bases de dados no jornalismo. E, seguindo a sua análise, consideramos que a aquisição e o uso das BDs no jornalismo podem ser explicados através do modelo da curva do “S”, utilizado dentro da teoria difusionista para observar a expansão das inovações tecnológicas. Neste sentido, incorporamos ao trabalho as contribuições desta teoria, por meio da obra de Everett Rogers (1995).

A expansão crescente da utilização das bases de dados e as suas diversas aplicações no jornalismo digital nos permite inferir a existência de algumas rupturas com determinados padrões nas esferas da gestão e da produção, da construção das narrativas ou peças informativas e na apresentação dos conteúdos jornalísticos. Neste aspecto, adotamos a idéia de ruptura tal qual trabalhada por Palacios (2002a, 2003b, 2003c; 2004), assim como a de potencializações, para compreender as mudanças que as BDs trazem para a criação e gestão de produtos jornalísticos digitais. Paralelamente, outras referências foram acrescentadas: Mielniczuk (2003); Viera (2004); Mielniczuk; Marques (2006); Schwingel (2003, 2004, 2005); Schwingel; Machado; Palacios; Rocha (2005); Silva Júnior, (2006); López (2006).

Como contraponto às rupturas, trazemos o conceito de *remediation*, formulado por Jay David Bolter & Richard Grusin (2000), para perceber também as continuidades e aperfeiçoamentos. Conforme os autores, as novas mídias digitais não são agentes externos que vêm transformar uma cultura insuspeitável. Ao contrário, elas emergem de dentro dos contextos culturais e refazem as outras mídias, as quais estão embutidas nos mesmos ou similares contextos. A internet, nesse caso, refaz todos os meios, melhorando-os em muitos

aspectos e acrescentando recursos novos, enquanto a web, especialmente, tem uma natureza remediadora, operando de modo híbrido e inclusivo (BOLTER & GRUSIN, 2000, p. 198).

Sendo o jornalismo digital uma modalidade que emerge com a internet, ele próprio remediará e será remediado por modalidades anteriores e por tecnologias como a das bases de dados, bem como irá originar inovações quanto aos modos de fazer jornalismo nas redes digitais, forjando, assim, um cenário de dupla via caracterizado por remediações e rupturas. Além das duas ações, surge uma terceira, a das potencialidades, que indicam panoramas possíveis em vias de constituição.

Para apreender os aspectos técnicos relacionados às bases de dados, metadados, *data mining*, algoritmos, suas funções, definições, aplicações, evoluções, compusemos um quadro integrado por estudos provenientes das ciências da computação e suas especialidades (DEEN, 1977; DATE, 1984; FRANGOMENI, 1987; ATTWOOD, MAIDENHEAD, 1989; MAURER, SCHERBAKOV, HALIM, RAZAK, 1998; PEREIRA, 1998; SOL, 1998; PIATTINI; CALERO, 1999; PIATTINI, 2000; ADAMS, 2001; BRICKLIN, 2001; SOUSA, 2002; VILLEGAS, GARCÍA, 2002; CARLEY, 2002; NAVEGA, 2002; BERLINSKI, 2002; GUIMARÃES, 2003; JORGENSEN, 2003; SALGADO, RODRÍGUEZ, 2003; CARA, BROADBENT, 2003; GRAY, 2004; PAUL, 2004; PORT DA ROCHA, 2004; DORNFEST, BRICKLEY, 2006), além de outros autores, cujos trabalhos trafegam entre a comunicação audiovisual e as ciências e sistemas de informação (BENÍTEZ, RODRÍGUEZ, 2005; QUIJADA, 2006).

Da abordagem teórica da tese constam, ainda, diversas referências relacionadas ao jornalismo e ao jornalismo digital (WOLF, 1995; BASTOS, 2000; ALVES, 2001, 2005; DÍAZ NOCI, 2001; CANAVILHAS, 2001, 2004, 2005, 2006; TRAQUINA, 2002; FONTCUBERTA, 2002; MOURIQUAND, 2002; BOCZKOWSKI, 2002, 2004; DE WOLK, 2001; EDO, 2002, 2005; DÍAZ NOCI, SALAVERRÍA, 2003; GRADIM, 2003, 2005;

BOWMAN, WILLIS, 2003; DEUZE, 2003, 2004; SALAVERRÍA, 2005a; SALAVERRÍA, 2005b; CABRERA, 2005; MCADAMS, 1995, 2005; HALL, 2005; FRANCISCATO, 2005; GRANIERI, 2005; ROSEN, 2003, 2005a, 2006a, 2006b; GILLMOR, 2004, 2005; RIBAS, 2005; NOGUEIRA, 2005; MOHERDAUI, 2005, PRYOR, 2006; FONTCUBERTA; BORRAT, 2006); à teoria literária: (MURRAY, 1997, 1999, 2003; RYAN, 2001), às novas mídias (BOLTER, 2001; LIEVROUW, LIVINGSTONE, 2002; BOLTER & GROMALA, 2003; LOVINK, 2003; MEADOWS, 2003; GOSCIOLA, 2003; GRAHAM, 2004; HERING, 2004; LIEVROUW, 2004), entre outras fontes sobre história e evolução dos meios de comunicação (LEVINSON, 1998; BRETON, PROULX, 2000; DIZARD JR, 2000; WOLTON, 2000; CASTELLS, 2001; BURKE, BRIGGS, 2004; PISCITELLI, 2005).

## V. Delimitação do objeto e casos estudados

O objeto de estudo desta tese é o **Jornalismo Digital em Base de Dados (JDBD)**. Situamos o surgimento deste modelo na fase intermediária entre a terceira (MIELNICZUK, 2003; PRYOR, 2002) e a quarta gerações de desenvolvimento desta modalidade, identificando funcionalidades e características do formato. A partir da análise dos produtos, definimos parâmetros que nos permitiram descrever particularidades e apontar as remediações, rupturas e potencialidades.

O *corpus* empírico é constituído por produtos jornalísticos digitais de perfis distintos: aqueles pertencentes a empresas informativas consolidadas, chamadas grandes *players* do *mainstream* da mídia; aqueles que são referência na rede, pelo caráter inovador e diferencial, mas não necessariamente pertencem a empresas jornalísticas; produtos experimentais, além daqueles laboratoriais criados e implementados em centros acadêmicos. Para a seleção dos

casos, consideraram-se alguns critérios, tais como: a) originalidade, b) representatividade e c) diversidade.

Optamos aqui por um conjunto de casos agregados, pois consideramos que, num contexto de remediações, rupturas e potencialidades, quanto mais amplo fosse o *corpus* empírico – refinado ao longo de toda a investigação à medida que realizávamos a operação articulada entre a observação sistemática e o estudo teórico –, poderíamos obter uma análise mais enriquecedora e abrangente sobre o nosso objeto, favorecendo, com isso, o processo de confirmação (ou não) das hipóteses indicadas acima. Nosso interesse voltou-se para produtos nos quais fosse possível identificar as características e as funcionalidades trabalhadas.

Como supúnhamos no período da análise preliminar e da observação livre para a seleção dos casos (2004), seria difícil encontrar num só produto todas as funcionalidades e categorias próprias desse modelo. Assim, optamos por realizar uma cartografia do objeto elegendo uma amostra mais ampla, constituída por exemplos representativos e de países distintos que nos permitissem demonstrar as especificidades que conformam o **JDBD**.

Temos, portanto, como produtos que integram o agregado de casos os que se seguem:

*Akademia*<sup>2</sup>, *BBC News*<sup>3</sup>, *blog Deu no jornal*<sup>4</sup>, *Elmundo.es*<sup>5</sup>, *Elpais.com*<sup>6</sup>, *Folha Online*<sup>7</sup>, *Google News*<sup>8</sup>, *International Herald Tribune*<sup>9</sup>, *Libération*<sup>10</sup>, *MSNBC*<sup>11</sup>, *News Brief*<sup>12</sup> e *News*

---

<sup>2</sup> Produto laboratorial do Labcom/UBI ([www.labcom.ubi.pt](http://www.labcom.ubi.pt)). Embora esteja atualmente desativado, ele é incorporado por ter sido um dos casos analisados no início do desenvolvimento da pesquisa. Parte do seu conteúdo foi integrada no outro produto da UBI: o jornal *Urbi et Orbi*.

<sup>3</sup> <http://news.bbc.co.uk>.

<sup>4</sup> <http://deunojornal.zip.net/>.

<sup>5</sup> [www.elmundo.es](http://www.elmundo.es).

<sup>6</sup> [www.elpais.com](http://www.elpais.com).

<sup>7</sup> <http://www.folha.uol.com.br/>.

<sup>8</sup> <http://news.google.com/?hl=en>.

<sup>9</sup> [www.iht.com](http://www.iht.com).

<sup>10</sup> <http://www.liberation.fr/>.

<sup>11</sup> [www.msnbc.msn.com](http://www.msnbc.msn.com).

<sup>12</sup> <http://press.jrc.it/NewsBrief/worldedition/en/en.html>.

*Explorer*<sup>13</sup> do *European Media Monitor (EMM)*, *Newzbubble*<sup>14</sup>, *Newsmap*<sup>15</sup>, *Notibits*<sup>16</sup>, *Panopticon*<sup>17</sup>, *Portal Terra*<sup>18</sup>, *TDC*<sup>19</sup>, *The New York Times*<sup>20</sup>, *Topix.net*<sup>21</sup>, *Urbi et Orbi*<sup>22</sup> e *Washingtonpost.com*<sup>23</sup>.

Certamente, esta é apenas uma amostra dentro de um universo muito maior de produtos jornalísticos digitais que estão empregando bases de dados. Contudo, nosso esforço aqui foi o de relacionar alguns dos casos representativos para fundamentar as discussões conceituais. Ao longo da tese, mencionamos também outros exemplos.

Pela natureza teórica da tese, os casos selecionados são empregados para ilustrar os argumentos, demonstrando a sua validade, testando ou refutando as hipóteses. Assim, as descobertas sobre os casos estão distribuídas nos capítulos que a compõem. Esse

---

<sup>13</sup> <http://press.jrc.it/NewsExplorer/dayedition/en/latest.html>.

<sup>14</sup> <http://www.newzbubble.com/index.php>.

<sup>15</sup> <http://www.marumushi.com/apps/newsmap/newsmap.cfm>.

<sup>16</sup> <http://www2.ing.puc.cl/~dcolle/notibits/index.htm>.

<sup>17</sup> <http://www.panopticon.ufba.br/>. Produto laboratorial da Faculdade de Comunicação da UFBA, que, atualmente, está em nova fase de implementação para incorporar a *Plataforma de Publicação e Ensino Panopticon*, um subproduto da Rede Latino-Americana para o Ensino do Jornalismo.

<sup>18</sup> [www.terra.com.br](http://www.terra.com.br).

<sup>19</sup> <http://www2.ing.puc.cl/~dcolle/notibits/index.htm>. Este é o link por meio do qual se pode ter acesso aos arquivos da *TDC (Temas de Tecnologías Digitales de Comunicación)*. O produto de caráter laboratorial-acadêmico esteve on-line entre 2000 e 2004 e foi um dos produtos analisados no início da pesquisa. A revista *Notibits* é o produto que de certa forma dá continuidade à *TDC*.

<sup>20</sup> [www.nytimes.com](http://www.nytimes.com).

<sup>21</sup> [www.topix.net](http://www.topix.net).

<sup>22</sup> [www.urbi.ubi.pt](http://www.urbi.ubi.pt).

<sup>23</sup> [www.washingtonpost.com](http://www.washingtonpost.com).

procedimento está de acordo com a metodologia de estudo de casos aplicada em outras investigações realizadas no âmbito do Grupo de Pesquisa em Jornalismo On-line, GJOL (MACHADO, PALACIOS, 2005). Sendo assim, evitou-se a descrição sistemática e pormenorizada dos produtos, e também uma discussão exaustiva dos resultados, pois que os dados empíricos são uma das fontes com que se opera.

## **VI. Metodologia**

Para alcançar os objetivos desta investigação, manejamos com bases teórico-conceituais multidisciplinares – do jornalismo às novas mídias, da teoria difusionista às ciências da computação. Este procedimento permitiu tratar o objeto sob o amparo de uma abordagem mais ampla e, com isso, estabelecer a fundamentação teórica. A pesquisa, de caráter exploratório (GIL, 2002), está ligada, ainda, ao tipo fundamental ou teórico, que tem por função aumentar a soma dos saberes disponíveis, criando quadros teóricos de referência (SANTAELLA, 2001, p.140).

Como modelo conceitual e operativo, adotamos o método do estudo de caso, considerado o mais apropriado para a investigação de fenômenos contemporâneos dentro do seu contexto real, conforme Robert Yin (1994, 2005) e Gregório Rodríguez (1996). Além disso, o estudo de caso está entre os mais usados na pesquisa em comunicação (LOPES, 2001, p.150), sendo o mais adequado para pesquisas exploratórias e particularmente útil para a geração de hipóteses.

Em vez do modelo convencional, empregamos o estudo de caso como ilustração (MACHADO, PALACIOS, 2005) para fundamentar as argumentações. Como já explicado, decidiu-se pela composição de um *corpus* empírico com um agregado de casos representativo, marcado pela diversidade de formatos e, na medida do possível, contemplando o caráter de

inovação dos produtos. Aliado ao método do estudo de caso, trabalhamos também dentro da perspectiva do método histórico, já que buscamos analisar o emprego de bases de dados no jornalismo através do tempo, para compreender a evolução do uso desta tecnologia até o momento atual, quando a identificamos com um novo *status* para a modalidade do jornalismo digital.

Percorremos quatro fases durante o desenvolvimento da pesquisa: a) levantamento bibliográfico, revisão de literatura, entrevistas semi-estruturadas, delimitação do objeto, formulação das hipóteses e objetivos, observação livre e análise preliminar dos produtos, exame de qualificação; b) estágio doutoral sanduíche, observação sistemática, composição do conjunto de casos, revisão e atualização bibliográfica, construção do conceito operacional, identificação de características e funcionalidades; c) observação sistemática, refinamento teórico-empírico e elaboração das categorias descritivas para análise e aplicação na observação estruturada dos produtos; d) processamento do material, redação e revisão final da tese. A tabela a seguir mostra as fases da pesquisa desenvolvida no período de maio de 2003 a janeiro de 2007.

<b>FASES / TEMPO</b>	<b>MAIO/2003 A DEZ/2004</b>	<b>JAN/2005 A DEZ/2005</b>	<b>JAN/2006 A AGO/2006</b>	<b>SET/2006 A JAN/2007</b>
<b>1ª FASE</b>				
<b>2ª FASE</b>				
<b>3ª FASE</b>				
<b>4ª FASE</b>				

Quadro 1 – As fases da pesquisa

A decisão pela observação livre de produtos de perfis distintos, na etapa inicial, nos habilitou a empreender o refinamento da amostra, com a seleção dos 22 *sites* jornalísticos e a observação sistemática, para, em seguida, realizarmos a observação estruturada aplicada para

verificar funcionalidades, bem como a adequação e a operacionalização das categorias descritivas. Dedicamos três a quatro horas de navegação diária nos *sites* europeus (comunidade europeia, portugueses, espanhóis, ingleses, franceses), norte-americanos, brasileiros, chilenos e africanos.

A observação dos *sites* foi guiada por um formulário (Apêndice II), empregado para verificar aspectos específicos contidos nos produtos que denotassem o emprego de bases de dados. De posse dos dados coletados, procedeu-se à análise deles, adotando-se um procedimento combinado de checagem teórico-empírica, com os devidos refinamentos que se fizeram necessários. Esta etapa nos permitiu o melhor reconhecimento das funcionalidades para que chegássemos à atribuição das categorias que conferem especificidade ao **Jornalismo Digital em Base de Dados (JDBD)**.

Ressaltamos que a esfera contemplada aqui é a do produto. Estamos conscientes, contudo, de que o acompanhamento das rotinas produtivas certamente nos permitiria preencher eventuais lacunas que a tese possa apresentar. Porém, nos limites deste trabalho, optou-se pela observação sistemática dos produtos, posto que seria inviável realizar estudo de campo em países distintos. Utilizamos, em adição, dados secundários, colhidos na bibliografia sobre jornalismo digital, jornalismo, entre outras fontes de áreas afins, de diferentes idiomas.

## **VII. Estrutura da tese**

O presente trabalho está dividido em cinco capítulos, além da introdução e das conclusões. Iniciamos o percurso com a contextualização sobre a propagação do uso da tecnologia de bases de dados na contemporaneidade, abordando o seu novo “lugar”, por meio da cartografia da genealogia para *A Forma Cultural Base de Dados*. O objetivo deste primeiro capítulo é explorar os relacionamentos entre bases de dados como uma tecnologia e sistema de informação e bases de dados como uma forma cultural, que constitui uma nova

maneira para estruturar a experiência humana, a nossa existência no mundo e o próprio mundo em si. Examina-se, então, alguns dos modos como as bases de dados se inserem na sociedade contemporânea: como modelo para produtos da nova mídia, como ferramenta de comunicação, como uma nova estética e uma hipernarrativa.

O capítulo 2, cujo título é *As Bases de Dados e o Jornalismo* traz o histórico sobre a presença desta tecnologia no campo específico do jornalismo. Para compreender a apropriação e as etapas desse processo evolutivo, nos baseamos na teoria da difusão de inovações tecnológicas e a sua curva do “S” para explicar o emprego e a aplicação das bases de dados no trabalho jornalístico. Localiza-se o uso inicial a partir da década de 70, apontando-se as utilizações posteriores durante os anos 80 e 90 até chegarmos ao ano 2000. Em nossa perspectiva, o emprego da tecnologia de BDs no jornalismo se encontra num renovado movimento ascendente da curva do “S”, experimentando nova fase de expansão, devido ao seu potencial para a gestão de produtos jornalísticos digitais. Só que, agora, o uso das BDs é pensado sob um novo referencial conceitual contemplando-as como um paradigma para o jornalismo digital na fase de transição da terceira para a quarta geração.

O título *O que é Jornalismo Digital em Base de Dados – “JDBD”* já indica o objetivo do capítulo 3, no qual nos concentramos a explicar como as BDs instauram um paradigma para o jornalismo digital no limiar de uma nova década de desenvolvimento para esta modalidade jornalística. De acordo com a nossa abordagem, o emprego da tecnologia de bases de dados é um dos aspectos-chave desse estágio, no qual se verifica o estabelecimento do paradigma *JDBD* e do seu potencial para fazer emergir uma nova metáfora para a apresentação dos conteúdos jornalísticos. A proposição de tal modelo é justificada, sobretudo, em função da complexificação dos processos para a implementação de produtos jornalísticos que vem sendo processada nos últimos três anos. Procedemos, assim, à definição e à caracterização desse estágio, discutindo os aspectos sobre a atribuição de um novo *status* para

as bases de dados no jornalismo digital, e trazendo o aporte teórico que corrobora a ampliação do significado das BDs. Em paralelo, buscamos ilustrar os conceitos através do material empírico que integra o conjunto de casos da amostra trabalhada. Fechamos esse terceiro capítulo destacando alguns pontos sobre os sistemas de gestão de conteúdos e com a descrição dos aspectos técnicos das BDs, segundo as formulações provenientes da área das ciências da computação.

No capítulo 4, *Delimitando o “JDBD” e as suas propriedades*, apresentamos o conceito para Jornalismo Digital em Base de Dados que colocamos em operação, identificando e explicando as 18 funcionalidades encontradas, bem como as sete categorias que conferem especificidade a este modelo. A demarcação das funcionalidades e das categorias propostas está amparada na observação do *corpus* empírico constituído por 22 *sites* jornalísticos de perfis distintos, associada à abordagem teórica construída. De acordo com o panorama que se delinea, consideramos as BDs como agentes fundamentais para gerar um jornalismo digital mais diferenciado, inteligente e dinâmico.

O capítulo 5, *Uma Nova Metáfora a Tomar Forma*, antecipa as conclusões e, assim, focaliza o que se coloca como os resultados prováveis da investigação realizada. Destacam-se, inicialmente, os pontos que evidenciam a emergência de uma metáfora particular para os produtos jornalísticos digitais, para, em seguida, explorar algumas das rupturas e potencialidades delineadas pelo modelo **JDBD**. Por último, no lugar reservado de fato às conclusões, revisamos os principais aspectos estudados durante a pesquisa, verificando o cumprimento dos objetivos e a confirmação (ou não) das hipóteses estabelecidas. Avalia-se, associadamente, o alcance dos resultados encontrados, ressaltando-se os pontos fortes e fracos do modelo proposto, as perspectivas, assim como as contribuições aportadas pela investigação realizada. Finaliza-se com a sugestão de continuidade para a pesquisa (em nível de pós-doutorado), para que se estude o modelo **JDBD** aplicado à esfera das rotinas produtivas.



## **A FORMA CULTURAL BASE DE DADOS**

## 1

**A forma cultural base de dados**

---

*Toda época é por definição única e possui a tecnologia a que pode ter acesso.*  
- Álvaro Vieira Pinto.

*(...) the key question, about technological response to a need,  
is less a question about the need itself than about its place in  
an existing social formation.*  
- Raymond Williams.

*If [with] the arrival of the Web the world appears to us as an  
endless and unstructured collection of images, texts, and other  
data records, it is only appropriate that we will be moved to model  
it as a database. But it is also appropriate that we would want to  
develop poetics, aesthetics and ethics of this database.*  
- Lev Manovich.

**1.1 O “LUGAR” DAS BASES DE DADOS NA CONTEMPORANEIDADE**

Criadas na década de 60 do século XX com o propósito de resolver problemas com sistemas orientados de arquivo, para os quais elas eram compactas, rápidas, seguras, fáceis de usar, precisas e permitiam o compartilhamento dos dados entre os usuários, as bases de dados<sup>24</sup> (BDs) despontavam como uma solução para a organização estruturada de vastos volumes de informações. Até aquele momento, depois das pessoas

---

<sup>24</sup> O primeiro uso conhecido do termo aconteceu em junho de 1963, nos Estados Unidos, aparecendo como duas palavras (*data base*), no título do simpósio *Development and Management of a Computer-centered Data Base*, patrocinado pela System Development Corporation. O termo *database* como uma única palavra se tornou comum no início da década de 70 na Europa.

e dos livros, eram as bibliotecas os locais onde se estocavam os numerosos repositórios de dados, fato que as tornaram conhecidas como as primeiras bases de dados reais.

A criação das primeiras bases de dados veio após a invenção dos grandes computadores e mudou a maneira de se armazenar, organizar, classificar, recuperar e compartilhar informações. À medida que a indústria da informática evoluía no aprimoramento do computador e no desenvolvimento de *softwares* e *hardwares*, implementações eram realizadas, recursos acrescentados e mais fácil se tornava o processo de estocagem, organização, recuperação e uso da informação, ao passo que também se expandiam as aplicações de BDs aos mais diversos setores nos vários campos do saber.

Como uma tecnologia e sistema de informação, as bases de dados vêm sendo empregadas há quatro décadas. Os aperfeiçoamentos e desenvolvimentos experimentados ao longo desse período as fizeram chegar à era da internet e da web como uma das mais imprescindíveis soluções para o armazenamento, a estruturação e o compartilhamento de informações, atendendo, assim, à dinâmica e às demandas e necessidades sociais. As BDs se tornaram coleções estruturadas de dados mantidas em computadores, como são mais popularmente compreendidas.

Selena Sol (1998), em sua *Introduction to databases for the web: Pt.1*, explica que, em uma definição simples, bases de dados são sistemas de registros computadorizados. E, numa definição mais completa, elas são um sistema que envolve dados, o *hardware*, que armazena fisicamente aquele dado; o *software*, que utiliza o sistema de arquivo do *hardware* para: 1) armazenar o dado e 2) prover um método

padronizado para recuperar ou mudar o dado; e, finalmente, os usuários, que vão transformar o dado<sup>25</sup> em informação<sup>26</sup>.

Baseada em Sol (1998), Christiane Paul (2004, p.02) vai apontar como um dos traços distintivos entre as bases de dados digitais e os seus predecessores analógicos a sua possibilidade inerente para a recuperação e filtragem do dado (ou dos registros nelas inseridos) em múltiplas maneiras. Ela define as BDs como:

A database essentially is a system that comprises the hardware that stores the data; the software that allows for housing the data in its respective container and for retrieving, filtering, and changing it, as well as the users who add a further level in understanding the data as information (PAUL, 2004, p.03)<sup>27</sup>.

Além desta, outras diferenças podem ser acrescentadas, como a descentralização, que permitiu passar das então centralizadas bases de dados para o padrão aberto constituído pela internet e pela web – ela própria considerada a maior das bases de dados atualmente disponível. O trabalho de Paul (2004) é trazido aqui, pois, indo além da definição desta tecnologia como um sistema, vai abordá-la como uma

---

<sup>25</sup> Dado, em geral, é entendido como fatos em sua forma primária. Em computação, dado é informação que foi traduzida em uma forma que é mais conveniente para ser processada. Dado é informação convertida na forma binária digital. In: *Whatis?com IT Encyclopedia and learning center*. Disponível em: [http://searchdatamanagement.techtarget.com/sDefinition/0,,sid91\\_gci211894,00.html](http://searchdatamanagement.techtarget.com/sDefinition/0,,sid91_gci211894,00.html).

<sup>26</sup> Estrutura significativa com a intenção de gerar conhecimento. Informação é estímulo que tem significado em algum contexto para o seu receptor. Quando a informação é inserida e armazenada em um computador, ela geralmente é referida como dado. Depois de processada, a saída do dado pode novamente ser percebida como informação. Quando a informação é empacotada ou usada para o entendimento ela é conhecida como conhecimento. In: *Whatis?com IT Encyclopedia and learning center*. Disponível em: [http://searchsqlserver.techtarget.com/sDefinition/0,,sid87\\_gci212343,00.html](http://searchsqlserver.techtarget.com/sDefinition/0,,sid87_gci212343,00.html).

<sup>27</sup> “Uma base de dados é essencialmente um sistema que constitui o hardware que armazena o dado: o software que permite alojar o dado em seu *container* respectivo e recuperá-lo, filtrá-lo e mudá-lo, tanto como os usuários que acrescentam um nível melhor no entendimento do dado como informação”. (t.A. tradução livre da Autora). Todas as citações em idioma estrangeiro contidas neste trabalho foram por nós traduzidas e aparecerão com a designação da sigla “t.A.”, tradução da Autora.

forma cultural. E, neste caso, nos importa diretamente, porque talvez resida nesta concepção o caráter mais distintivo para as bases de dados na contemporaneidade.

Coube a Lev Manovich, no livro *The Language of New Media* (2001), reconhecer as bases de dados como uma forma cultural simbólica na por ele denominada *computer era* ou era do computador. Ele foi um dos primeiros a sistematizar elaborações teóricas acerca da ubiqüidade das bases de dados, percebendo-as como lógica e estrutura para forjar boa parte dos objetos da nova mídia. Neste sentido, para Manovich, o entendimento de forma cultural está relacionado com modalidades tecnológicas através das quais pode ser estabelecida uma relação homem-conteúdo.

A base de dados seria a primeira forma cultural contemporânea, ao passo que o espaço navegável, virtual e interativo em 3D – empregado em jogos de computador, animação e nas Interfaces Homem-Computador (HCI) – seria a segunda forma-chave (MANOVICH, 2001, p.252). Ele associa a primeira forma com trabalho (o trabalho pós-industrial de processamento de informação) e a segunda com o lazer e a diversão, exemplificada pelos *games* (MANOVICH, 2001, p.215).

A concepção de bases de dados como forma cultural tem como referência original as análises elaboradas por Ervin Panofsky, teórico da história da arte, sobre a perspectiva linear como uma forma simbólica da era moderna. Seguindo Panofsky, Manovich vai pensar as bases de dados como a nova forma simbólica da “sociedade computadorizada”<sup>28</sup> – um novo modo para estruturar a experiência humana, a nossa existência no mundo e o próprio mundo em si (MANOVICH, 2001, p.219).

Surgida com a World Wide Web, a nova mídia nos leva a experienciar uma nova maneira de perceber o mundo. Na definição do autor, a nova mídia surge a partir da

---

<sup>28</sup> Termo aplicado por Jean-François Lyotard no livro *The Postmodern Condition*, lançado em 1979.

síntese entre a computação e a tecnologia da mídia e tem o computador como principal instrumento afetando todos os estágios da comunicação: aquisição, manipulação, armazenamento, distribuição e convergência. O resultado advindo da utilização do computador é a mudança de toda a cultura para formas de produção mediadas por esta máquina. Quanto aos objetos da nova mídia, eles tanto podem ser novos como os já existentes, que têm sua forma afetada pelo uso do computador.

É com as implementações e aperfeiçoamentos providos pela computação que a nova mídia se desenvolve e tem estabelecido a sua identidade, principalmente, através de termos tecnológicos como representação discreta e numérica, escalabilidade, automação, variabilidade, e transcodificação – princípios da nova mídia detalhados mais abaixo no subtópico “1.1.2 Web como base de dados”. No seu processo de investigação para a formulação de uma teoria para a nova mídia, Lev Manovich aponta como necessário voltar-se para as ciências da computação, para compreender a sua lógica e encontrar termos novos, categorias e operações que caracterizam a mídia que se torna programável, e bases de dados ou *database* é um deles (MANOVICH, 2001, p. 48).

So, for instance, a computer database is quite different from a traditional collection of documents: it allows one to quickly access, sort, and recognize millions of records; it can contain different media types, and it assumes multiple indexing of data, since each record besides the data itself contain a number of fields with user-defined values (MANOVICH, 2001, p.214)<sup>29</sup>.

Nesta acepção, as bases de dados podem armazenar individualmente ou a partir de diversas combinações qualquer tipo de dado ou objeto digital, desde documentos

---

<sup>29</sup> “Deste modo, em primeira instância, uma base de dados computadorizada é totalmente diferente de uma coleção tradicional de documentos: ela permite a alguém acessar rapidamente, classificar e reconhecer milhões de registros; ela pode conter diferentes tipos de mídia e supõe múltipla indexação de dados, desde que cada registro ao lado do próprio dado contenha um número de campos com valores usuário definidos”. (t.A.)

textuais, a fotografias, clipes, seqüências de áudio, imagens estáticas, em movimento, ou, ainda, animações, mapas, gráficos, entre outros, que podem ser navegados/ acessados de modos variados. Note-se, ainda, que, ao argumentar em favor da BD como forma cultural simbólica, Lev Manovich a apresenta como aquela que está para a era do computador, assim como a novela e o cinema estiveram para a era moderna. E, como forma do contemporâneo, a BD representa o mundo como uma coleção de itens, sobre a qual o usuário performará várias operações, como: ver, navegar, buscar (MANOVICH, 2001, p.219).

Tal lógica está implícita na estruturação de boa parte dos produtos da nova mídia, como o CD-ROM de um museu virtual com sua coleção de imagens para serem acessadas de diferentes maneiras, e um *web site*, que apresenta uma lista seqüencial de elementos separados: blocos de texto, imagens, videoclipes e *links* para outros *sites*. Logo, na perspectiva construída pelo autor, as bases de dados se tornam o centro do processo criativo do *design* da nova mídia.

O entendimento do potencial das bases de dados na era do computador, portanto, amplifica o sentido do termo, indo além daquela noção mais básica de estrutura de dados para rápida recuperação, que até então norteou os procedimentos de armazenamento e ordenamento de informações, para adquirir o *status* de nova forma cultural simbólica. Neste sentido, Manovich declara:

Indeed, if after the death of God (Nietzsche), the end of grand Narratives of Enlightenment (Lyotard), and the arrival of the Web (Tim-Berners Lee), the world appears to us as an endless and unstructured collection of images, texts, and the other data records, it is only appropriate that we will be moved to model it as a database.

But is also appropriate that we would want to develop a poetics, aesthetics, and ethics of this database (MANOVICH, 2001, p.219)<sup>30</sup>.

Cabe acrescentar, ainda, o sentido metafórico com que o autor usa o termo base de dados, isto é, ele não o aplica apenas estritamente para bases de dados, mas em sentido mais amplo. Provavelmente, este *status* atribuído por Lev Manovich às bases de dados como uma forma-chave tenha encontrado mais repercussão entre artistas da nova mídia (assunto do subtópico “1.2.1 Uma estética: modos diferenciados para visualização da informação”) do que mesmo entre seus pares na academia<sup>31</sup>, mais interessados em focalizar outros aspectos das formulações teóricas de Manovich, como, por exemplo, a escolha da abordagem da nova mídia tendo como parâmetro o cinema e os princípios que a caracterizam. No entanto, a sua proposição<sup>32</sup> nos permite perceber como a propagação do uso desta tecnologia a coloca como elemento central na conformação de boa parte dos produtos culturais ora em circulação e, por extensão, na configuração da própria sociedade.

---

<sup>30</sup> “Na verdade, se depois da morte de Deus (Nietzche), do fim das imponentes Narrativas do Iluminismo (Lyotard) e da chegada da Web (Tim-Berners Lee), o mundo aparece para nós como uma infinita e desestruturada coleção de imagens, textos e outros registros de dados, é apropriado que nós estejamos mudando para modelá-lo como uma base de dados. Mas é também apropriado que nós vamos querer desenvolver uma poética, estética e ética destas bases de dados”. (t.A.).

<sup>31</sup> Enquanto algumas revisões sobre o seu livro destacam o caráter inovador e provocador das formulações teóricas do autor, outras, por outro lado, embora o considerem valioso, talvez tentem minimizá-lo, ao classificarem como divertido, uma fonte de entretenimento sobre a nova mídia. Ver, por exemplo: *Computable Culture and the Closure of the Media Paradigm: Lev Manovich's The Language of New Media*, de William B. Warner (s/d), disponível em: [http://www.manovich.net/LNM/review\\_Warner.doc](http://www.manovich.net/LNM/review_Warner.doc); e *Metonymical Mov(i)es*, de Ink Arns (2002), disponível em: <http://www.projects.v2.nl/~arns/Texts/Media/manovich-review.html> .

<sup>32</sup> A idéia de bases de dados como forma simbólica foi apresentada pelo autor no artigo *Database as a Genre of New Media*, publicado no *AI Society: The Journal of Human Centred and Machine Intelligence*, em 1999. Segundo Manovich, as idéias contidas no artigo haviam sido apresentadas no simpósio Chips and Bits, realizado na UCLA School of Film and Television, em 1997. Numa versão posterior, o artigo teve como título *Database as Symbolic Form*, sendo publicado no livro (2001) como um dos capítulos, o 5, no qual ele discute as formas da nova mídia.

Mesmo presente na vida diária, certamente, a importância dada ao termo base de dados ainda é reduzida. Considerando que, diariamente, bilhões e bilhões de registros são armazenados em bases de dados compartilhadas entre inúmeras instituições e órgãos governamentais, empresas e entidades, em geral, escapa a uma grande maioria a compreensão sobre o valor incomensurável das informações produzidas sobre nós mesmos circulando nos milhares de bases de dados existentes atualmente. Nessas redes virtuais, somos registros de bases de dados. São elas que vão constituir e multiplicar nossas identidades, armazenando-as e transmitindo-as por entre essas redes<sup>33</sup>. No cenário contemporâneo no qual corre a vida quotidiana, o mundo cada vez mais é visto como uma base de dados – um depósito de descrições, números, fatos (NOGUEIRA, 2001).

As BDs deslocam linhas de força e passam a exercer papel fundamental, seja registrando, diariamente, bilhões de ações e hábitos de todos nós, produzindo perfis, formatando novos padrões de consumo – ou a personalização em massa –, gerando um novo “sujeito”, forjando produtos culturais, armazenando, estruturando e distribuindo conhecimento. Ou seja, um mundo traduzido em bases de dados. Acreditamos que este novo “lugar” para as bases de dados na contemporaneidade corrobora a concepção de forma cultural.

Na cartografia do fenômeno, encontramos formulações que vão caracterizar as bases de dados como um modelo para produtos da nova mídia (MANOVICH, 2001), como uma estética e hipernarrativa (VESNA, 1999, 2000a, 2000b; MANOVICH, 1999,

---

<sup>33</sup> Esse mundo-base de dados traz, por outro lado, conseqüências negativas para a privacidade dos indivíduos. De acordo com autores como Mark Poster (1995, 2000), por exemplo, as BDs, com a capacidade que possuem para constituir e multiplicar a identidade do indivíduo, vão também promover o seu controle. Poster (1995, p.64) defende que, com as bases de dados informáticas, um novo discurso, uma nova prática opera no campo social – um *super-panopticon* –, que reconfigura o modo de constituição do sujeito. Ao contrário do *panopticon*, os “detidos” não necessitam ficar albergados num tipo qualquer de arquitetura; apenas necessitam prosseguir com a sua vida quotidiana. Portanto, diz Poster (1995, p.69), o *super-panopticon* é menos obstaculizante que o seu predecessor, no entanto, não é menos eficiente na sua tarefa de normalização.

2000, 2001, 2006; DIETZ, 2000, 2004; DIETRICH, 2000; CISLER, 2000; LEÃO, 2003; PAUL, 2004), como uma ferramenta de comunicação de grande valor para o prognóstico de cenários futuros e de relevância crescente para a tomada de decisões (LUCAS, 2004, 2006), como um formato (MACHADO, 2004a, 2004b, 2004c, 2006) e como um paradigma e como uma metáfora (BARBOSA, 2004a, 2004b, 2004c, 2005a, 2005b, 2006a, 2006b) no jornalismo digital.

Tais formulações nos fornecem análises que corroboram a idéia do alargamento do significado das bases de dados na contemporaneidade, como também demonstram os usos, aplicações, vantagens, especificidades e conseqüências do emprego expandido desta tecnologia. Indicam, notadamente, uma ruptura com o entendimento mais comum sobre bases de dados, até então circunscritas ao campo específico das ciências da computação.

Neste capítulo, nosso objetivo é explorar os relacionamentos entre bases de dados como uma tecnologia e sistema de informação e bases de dados como uma forma cultural. Cartografar a genealogia de uma nova forma e os modos como ela se apresenta na sociedade contemporânea: modelo para produtos da nova mídia, ferramenta de comunicação, nova estética, hipernarrativa, formato, metáfora e paradigma no jornalismo digital.

### **1.1.1 Como uma tecnologia passa a “forma cultural”**

O surgimento de uma tecnologia depende de um complexo contínuo de interação entre novas necessidades sociais, invenções, desenvolvimento e aplicação. Não surge como um evento único isolado ou série de eventos, mas diretamente conectada com um

processo de descobertas, que vão abrindo caminho para os melhoramentos técnicos, conforme também evoluem as sociedades. Várias tecnologias em momentos diferenciados de transformações sociais gerais surgiram como resultado desse complexo contínuo, como classificado por Raymond Williams (1990), entre eles, o telégrafo, o rádio, a televisão e, poderíamos acrescentar, também a internet e a WWW.

Em comum, elas possuem a característica de terem sido empregadas, inicialmente, como um aparato militar, para, em seguida, após períodos de aperfeiçoamentos e mesmo de necessidades variadas, controvérsias e até pressões econômicas e políticas, serem apropriadas socialmente. Esta apropriação, em geral, se dá em ritmos diferenciados e segundo circunstâncias particulares, próprias de cada país. O ponto central neste processo reside na incorporação da tecnologia como prática. Ou seja, quando ela é incorporada nos vários níveis da vida social e dos campos do saber, estágio no qual ela também alcança o potencial para fazer brotar novas formas culturais.

Com Williams (1990), podemos perceber o caminho que uma tecnologia percorre até se transformar em uma forma cultural. O contexto da sua análise é o início dos anos 70 e se reporta ao desenvolvimento de uma nova tecnologia de comunicação social – a televisão – que, àquela altura, inaugurava uma nova forma cultural, que passaria a conviver com outras formas universais já estabelecidas, como o jornal, o livro e o cinema. Ele foi um dos primeiros a considerar o aspecto cultural como fundamental no processo de apropriação de uma tecnologia. O estudo referencial *Television. Technology and Cultural Form* (1990), publicado pela primeira vez em 1974, possui muitos méritos, entre eles, o de ter trazido a questão cultural para o centro do debate sobre a tecnologia e o conceito de fluxo como organizador da programação. Para a discussão que aqui estabelecemos, as contribuições da obra de Raymond Williams nos

servem para entender como o surgimento de formas culturais também está relacionado com o uso de determinada tecnologia.

Ampliando o espectro para a relação de apropriação da técnica pelo homem, os escritos produzidos no início da década de 70 pelo filósofo brasileiro Álvaro Vieira Pinto – que resultaram na obra *O conceito de tecnologia* (2005) – nos mostram que a história da técnica, e, em sentido correlato, a das tecnologias, tem de ser evidentemente a história das produções humanas, integralmente entendidas, isto é, no estágio superior, social da atividade do produtor (PINTO, 2005, p.64). Vieira Pinto, para quem a expressão “era tecnológica” refere-se a qualquer época da história e não apenas aos tempos atuais, evidencia que:

São as técnicas, enquanto ações humanas concretas, que têm valor primordial porque se referem à relação direta de caráter problemático, do homem com o mundo, ao passo que a cultura designa apenas o conjunto delas em determinado tempo e lugar, mais as crenças e valores a elas agregados (PINTO, 2005, p.65).

Prosseguindo, esclarece que as técnicas ou tecnologias disponíveis em determinada fase do desenvolvimento humano configuram certos conceitos e não o inverso. Se a tecnologia de cada fase histórica permite ampliar a capacidade humana de domínio da natureza, com as respectivas elaborações teóricas, isso ocorrerá, como ele diz, até certo ponto. Pois o processo evolutivo é dinâmico e, quando se esgotam as possibilidades de revelação de novos aspectos do universo, a capacidade humana de apreensão da realidade e de invenção de instrumentos de ação, materiais ou teóricos, não estaciona. Ao contrário, impulsiona novo conhecimento, nova fase de exploração de possibilidades, invenções, procedimentos e explicações teóricas, que vão gerar novo

estágio de evolução tecnológica, e, com isso, um novo estágio para a criação da produção material, artística e filosófica.

Em relação às quatro acepções do termo “tecnologia” apresentadas por Álvaro Vieira Pinto (2005), acreditamos ser a primeira delas a que mais se afina com as formulações aqui propostas, pois indica que a tecnologia tem que ser a teoria, a ciência, o estudo, a discussão da técnica, abrangidos nesta última noção as artes, as habilidades do fazer, as profissões e os modos de produzir alguma coisa (VIEIRA PINTO, 2005, p.219). Não à toa, Vieira Pinto assinala ser esta a acepção primordial sobre tecnologia, pois abre à compreensão das demais<sup>34</sup>.

Retomando Williams (1990), encontramos que as várias necessidades surgidas em períodos distintos e os estilos de vida que caracterizam as sociedades no correr da história apontam que as formas são dinâmicas e os modos como se processam o seu reconhecimento são não apenas possíveis, como necessários. Assim como as formas encontram necessidades sociais de um novo tipo para ser estabelecidas, também serão necessárias novas definições sociais. Ao lado disso, podemos apreender, ainda, que formas culturais diferenciadas convivem ao mesmo tempo e vão sendo apropriadas de acordo com as particularidades de cada campo social, ainda que possam ser específicas de um campo particular.

Com as bases de dados, podemos inferir que ocorre algo desta natureza. Como uma tecnologia e sistema de informação já em uso desde os anos 60, e sendo uma forma específica do domínio das ciências da computação para solucionar os problemas de

---

<sup>34</sup> As outras três acepções são: tecnologia como equivalente pura e simplesmente da técnica (sentido mais popular da palavra); tecnologia como o conjunto de todas as técnicas de que dispõe uma determinada sociedade, em qualquer fase de seu desenvolvimento; e tecnologia como ideologização da técnica, significado para o qual o filósofo dedica maior atenção no seu ensaio por ter, segundo ele, importância capital, haja vista a sua proximidade com o pensamento marxista (PINTO, 2005, p.219-220).

organização, armazenamento, recuperação e compartilhamento de informações, a flexibilidade das BDs, aliada aos desenvolvimentos das redes interativas digitais, do computador, dos *softwares* e das linguagens de programação, as coloca num patamar diferenciado, revestindo-as de novo significado: de forma cultural, delineada nos anos finais da década de 90 e em consolidação nesta primeira década do século XXI.

E para quê? Para dar conta de necessidades sociais de tipos novos num contexto de sociedade contemporânea complexa, na qual as tecnologias de informação e comunicação permeiam as atividades da vida diária, seja no ambiente de trabalho, na educação e na produção do conhecimento, nos modos de informar, criar e consumir informações, nas esferas da criação artística e cultural, do lazer e do entretenimento, da indústria, da economia e dos serviços, das relações sociais, enfim.

É com a internet – ela mesma um ambiente para estabelecimento de diversas formas culturais, capazes de constituir relações entre os diversos atores e criar novas convenções – que a forma base de dados floresce. É o que assinala Lev Manovich (2001), observando que os trabalhos multimídia com conteúdo “cultural” favorecem particularmente a forma base de dados. “O domínio da forma base de dados tem como exemplos mais óbvios as enciclopédias, coleções por definição, tanto como o CD-ROM (ou DVD), que caracteriza coleções de receitas, citações, fotografias, entre outros”<sup>35</sup> (MANOVICH, 2001, p.219).

Apesar de o autor não se referir à obra de Raymond Williams (1990) em sua construção teórica sobre a nova mídia, julgamos pertinente assinalar essa articulação entre as concepções de forma cultural trabalhadas por ambos. De alguma maneira, elas

---

<sup>35</sup> “The most obvious examples are popular multimedia encyclopedias, collections by definition, as well as other commercial CD-ROM (or DVD), that feature collections of recipes, quotations, photographs, and so on”. (t.A.).

convergem, na medida em que consideram a tecnologia como fator propulsor para o estabelecimento de formas culturais, em função da apropriação e representação sociais.

Williams focalizou a televisão para demonstrar como uma tecnologia de comunicação social legava uma forma cultural, enquanto Manovich nomeia as bases de dados como principal forma contemporânea, ao lado do espaço navegável.

Lev Manovich optou por uma análise que privilegia o cinema como modelo para estabelecer o paralelo entre as formas culturais contemporâneas, ou da nova mídia, e as formas precedentes (imprensa, fotografia, televisão). Entre os motivos, o fato de que o cinema se consolida enquanto forma cultural em uma virada de século (do XIX para o XX), assim como as formas atuais, que despontam entre o século XX e o XXI, e também porque o cinema opera dentro de uma lógica de síntese de modelos narrativos prévios. Ao eleger o cinema, ele quis mostrar como as novas mídias se apropriam de formas e convenções prévias, enfatizando, deste modo, as continuidades entre o novo e o antigo, e as interações entre a repetição histórica e a inovação (MANOVICH, 2001, p. 285).

Em sua perspectiva, o cinema ainda condiciona, um século depois, os elementos presentes nos modelos apoiados em tecnologias digitais. O cinema seria, deste modo, uma forma cultural consolidada, que sintetiza linguagens prévias e, por outro lado, determina uma nova síntese, que se processa entre os dispositivos tecnológicos digitais e as formas culturais.

O computador digital, por seu turno, torna-se o “meio perfeito para a forma base de dados” (MANOVICH, 2001, p.234). Esta, por conseguinte, servirá de estrutura e lógica para a criação de boa parte dos produtos culturais, sejam eles novos ou já existentes, que terão sua forma afetada pelo uso do computador. Com isso, as BDs se transformam em complexos de armazenagem de formas culturais. E, como uma última

observação importante, aportada na investigação de Elias Machado (2004b, 2006), como forma cultural particular, as bases de dados vão viabilizar, ainda, e pela primeira vez, que o espaço seja alçado à categoria de suporte, tal como o papel, a fotografia, o áudio e o vídeo (BOLTER, 1991).

### 1.1.2 Web como base de dados

É a face multimídia e interativa da internet – a World Wide Web – que opera transformando todo *site* em um tipo de base de dados. Na sua estrutura definida pela linguagem de formatação HTML, uma lista seqüencial de elementos separados (texto, fotografias, imagens em movimento, infografias, arquivos de áudio, entre outros) permite que se acrescentem novos componentes e *links*, o que faz com que os *sites* estejam sempre crescendo toda vez que se adiciona algo novo. Se novos elementos estão sendo acrescentados continuamente, os *web sites* dificilmente atingirão o estado de completude total. Mais coerente seria dizer, pois, que o resultado aponta mais para serem constituídos como coleções.

Uma vez digitalizados, os elementos ou dados podem ser organizados e indexados a partir de inúmeras possibilidades combinatórias. Aqui, podemos estabelecer relação com um dos seis princípios abstratos do hipertexto propostos por Pierre Lévy (1993, p.25-26): o princípio da exterioridade<sup>36</sup>, fundamental para preservar o caráter aberto da rede.

---

<sup>36</sup> Os outros princípios pensados por Lévy para preservar as possibilidades de múltiplas interpretações do modelo de hipertexto são: o da metamorfose, heterogeneidade, multiplicidade de encaixe das escalas, topologia e o da mobilidade dos centros (LÉVY, 1993, p.25-26).

De acordo com Lévy (1993), o crescimento e a diminuição de uma rede (e nisto podemos considerar o próprio *site* enquanto microrrede), sua composição, alimentação e recomposição permanente dependem de um exterior indeterminado: adição de novos elementos, conexões com outras redes (ou conexões com outros *sites* por meio dos *links* ou mesmo remetimentos a vários documentos). Tal princípio reforça, por outro lado, o quanto a descentralização da produção de conteúdos pode funcionar como um agente ativador da rede, uma vez que assegura a sua permanente alimentação, sendo operada de maneira contínua, em fluxo.

O que expomos acima também é explicado por meio dos cinco princípios sistematizados por Lev Manovich (2001, p.19-48) para demarcar algumas das principais diferenças entre a antiga<sup>37</sup> e a nova mídia. Recomendados para ser considerados muito mais como tendências gerais de uma cultura computadorizada do que como leis absolutas, esses princípios traçam um perfil da identidade da nova mídia. São eles:

- Representação numérica. O código digital permite trabalhar ou modificar cada objeto da nova mídia automaticamente, pois é programável;
- Modularidade. Um objeto da nova mídia tem a mesma estrutura em diferentes escalas tal qual um fractal<sup>38</sup>, o que significa que elementos de mídia, como texto, som, imagem, podem estar reunidos numa única

---

<sup>37</sup> A imprensa, a fotografia e a televisão.

<sup>38</sup> Um fractal – termo cunhado em 1975 por Benoit Mandelbrot para descrever um objeto geométrico que nunca perde a sua estrutura qualquer que seja a distância de visão – é uma forma geométrica, de aspecto irregular ou fragmentado, que pode ser subdividida indefinidamente em partes, as quais, de certa maneira, são cópias reduzidas do todo. A palavra fractal significa, sobretudo, auto-semelhante. Auto-semelhança é a simetria através das escalas. Ou seja, um objeto que possui auto-semelhança apresenta sempre o mesmo aspecto a qualquer escala que seja observado. Troncos de árvore, nuvens, montanhas são objetos que podem ser representados por fractais. Em 1980, Mandelbrot descobriu o primeiro fractal gerado por computador. A geometria fractal que usa softwares sofisticados produz imagens belas e interessantes, mixando arte e matemática.

narrativa ou documento, mas continuam mantendo suas identidades separadas. Exemplo disso são os *sites* que formam a web. Eles são constituídos por diferentes elementos de mídia e cada um deles pode ser acessado e modificado separadamente. Daí o motivo pelo qual Manovich considera a web completamente modular;

- Automação. O código numérico da mídia (princípio 1) e a estrutura modular de um objeto da mídia (princípio 2) permitem a automação de muitas operações, desde a criação, a manipulação até o acesso da mídia. A automação também traz a necessidade do acesso a bases de dados e pressupõe que o usuário possa criar ou modificar um objeto de mídia, vide técnicas incluídas em softwares comerciais, como processadores de texto, editores de *sites*, *Photoshop*, entre outros. Ou, num nível mais avançado, a automação por meio da criação de jogos que utilizam recursos da inteligência artificial e requer do computador entendimento dos significados inseridos nos objetos gerados;
- Variabilidade. Prevê que um objeto da mídia existe, potencialmente, em diferentes e infinitas versões. Como exemplo, uma fotografia. Ela pode ser usada enquanto ilustração de um texto, integrar uma galeria de fotos de tema específico, um *slideshow* e, ainda, pode ser disponibilizada juntamente com uma narração em áudio associada, na qual o fotógrafo descreve o processo de captura daquela determinada imagem. Ou seja, tem-se tanto uma variabilidade de modalidade como também de formato;
- Transcodificação. Este, talvez, seja o mais importante entre os princípios propostos. Consiste em traduzir uma forma cultural em outro

formato. Ou seja, todos os objetos da nova mídia podem ser traduzidos para outros formatos. O aspecto essencial é que a estrutura da mídia digital segue convenções estabelecidas na organização de dados pelo computador e essa estrutura é que permite o trânsito e compatibilidade transcodificável para outros arquivos de computador. O autor aponta que a lógica operacional do computador pode ser encarada como uma camada que afeta a camada cultural (integrada por categorias como enciclopédia e histórias curtas, narrativa e enredo, composição e ponto de vista, comédia e tragédia). Já a camada do computador é formada por *process* e *pocket*, a exemplo dos pacotes de dados transmitidos através da rede, funções e variáveis, linguagens do computador e estrutura de dados.

Assim, vemos a internet e a web como ambientes para estabelecimento de práticas e de formas culturais. Conforme Palacios (2003a), a internet é um ambiente compartilhado de comunicação, informação e ação para uma multiplicidade de (sub) sistemas sociais (economia, arte, jornalismo) e para agentes cognitivos (humanos). Enquanto para Manuel Castells (2001), a “Galáxia Internet” é um novo entorno de comunicação, uma nova estrutura social, que se está estabelecendo em todo o planeta para a vida das pessoas, segundo sua história, cultura e instituições (CASTELLS, 2001, p.305).

Sendo a internet também um grande repositório de informação, ela fornece as condições para o surgimento de novas formas culturais. Formas essas que, em boa medida, usarão as BDs como estrutura e suporte para indefinidos tipos de *sites* que circularão na web.

## 1.2 INTERFACE PARA PRODUTOS DA NOVA MÍDIA

Com o computador, vão dizer Bolter e Gromala (2003, p.168), a tarefa do empoderamento da interface passa a ser compartilhada entre o escritor/autor, o leitor/usuário e o próprio sistema de computação. Esta idéia da Interface Homem-Computador (*Human-Computer Interface, HCI*) foi concebida quando o termo computador era empregado somente como uma ferramenta primária. Contudo, durante a década de 90, a identidade dele mudou, e o computador se transformou numa máquina de mídia universal, usada não apenas para a autoria ou a composição de textos, mas também para armazenar, distribuir e acessar todas as mídias (MANOVICH, 2001, p.69). Uma interface é o todo da tela que se apresenta diante do usuário e é denominada habitualmente como Interface Gráfica para o Usuário (*Graphical User Interface, GUI*).

Como a distribuição de todas as formas de cultura se tornou baseada no computador, a interface é a forma pela qual todas as produções artísticas e culturais passam a ser mediadas, conforme assinalam em suas investigações Steven Johnson (2001) e Lev Manovich (2001). Desta maneira, emprega-se o termo “interface cultural” para descrever uma interface homem-computador-cultura, isto é, os modos através dos quais computadores nos habilitam a interagir com o dado cultural (leia-se textos, fotografias, filmes, música, ambientes virtuais). Ou, mais precisamente, um webjornal, um jogo de computador ou um *site* de uma banda de rock (PALACIOS, 2005).

Interessa notar que a interface molda no usuário a concepção de conceitos de funcionamento do computador e também da estruturação de mídias acessadas através

dele. E, nesse aspecto, mais do que se tornar transparente ou neutra<sup>39</sup>, a interface se torna ainda mais visível, e o usuário mais consciente da sua presença, pelas relações de mediação que engendra e pela interação que estabelece entre a máquina e o usuário e deste com os conteúdos.

Interfaces culturais incluem as interfaces utilizadas pelos *designers* de *web sites*, CD-ROMs, títulos de DVD, enciclopédias multimídia, museus, revistas on-line, webjornais, *sites* noticiosos em geral, portais, *weblogs*, e outros tantos objetos culturais da nova mídia, a exemplo dos projetos de *webart*. Como as bases de dados estão no centro do processo criativo do *design* da maioria dos produtos da nova mídia, elas vão constituir-se como uma estrutura que contém elementos, ou registros, e estabelecem relações entre eles traduzidas pela organização das interfaces e pela ordem de acesso do usuário (MANOVICH, 2001).

Mais enfaticamente falando, os objetos da nova mídia consistem de uma ou mais interfaces para uma base de dados de material multimídia. Isso ocorre porque o conteúdo do trabalho é separado da interface. Logo, torna possível criar diferentes versões para o mesmo material (princípio da variabilidade), como, por exemplo, no caso de uma fotografia num *site* jornalístico: ela tanto pode aparecer ilustrando uma notícia, como pode integrar uma galeria, um álbum de fotos, um *slideshow* e, ainda, ser disponibilizada junto com o áudio de um repórter relatando um determinado acontecimento. Em cada um desses casos, uma interface diferenciada é apresentada para que o usuário acesse/explore aquele material.

No campo da comunicação, Elias Machado (2004b, 2006) argumenta que mais do que uma estrutura lógico-matemática, que possibilita a organização, armazenamento

---

<sup>39</sup> Há uma discussão estabelecida sobre os aspectos inerentes à interface. Bolter e Grusin (2000), Bolter e Gromala (2003), Manovich (2001), Scolari (2004) são alguns dos que abordam o assunto. Aqui, não nos propomos a discutir a transparência, materialidade ou reflexividade da interface, mas julgamos pertinente destacar como ela é sempre mais evidenciada no processo de acesso e exploração de produtos culturais.

e recuperação de informações individualizadas, a base de dados aparece para os usuários como uma interface tipificada no espaço navegável, permitindo explorar, compor, recuperar e interagir com as narrativas.

Os professores da Universidade de Santiago de Compostela (Espanha) Xosé López, Manuel Gago e Xosé Pereira (2003, 2005, 2006) consideram as bases de dados como elementos essenciais que armazenam toda a informação, incluindo a estrutura lógico-matemática para a organização dos dados e a arquitetura da informação<sup>40</sup> que orienta a consulta, a recuperação da informação e possibilita a composição das narrativas acessadas por meio da interface gráfica. Os investigadores observam que tanto os jornalistas como os usuários sempre que inserem uma notícia (tarefa dos primeiros) ou consultam e extraem informação (tarefa dos segundos) em seu *site* favorito, na realidade, estão consultando e explorando as informações contidas em uma base de dados. Porém, como os usuários também podem inserir conteúdos, o mais correto seria incluí-los, igualmente, como produtores de informação nesse processo, de acordo com Machado (2004b, 2006)<sup>41</sup>.

A seguir, discutem-se aspectos relacionados aos modos diferenciados de visualização dos dados estruturados na forma base de dados, plasmados na chamada estética base de dados, bem como as formulações que apontam para a redefinição da

---

<sup>40</sup> Conceito criado por Richard Wurman em 1962, que o propôs como um mapa ou estrutura de informação que permite aos usuários encontrar seu caminho pessoal ao conhecimento. O profissional encarregado de criar esse mapa é o arquiteto da informação. Em 1998, Louis Rosenfeld e Peter Morville atualizaram o conceito, incorporando a recuperação da informação e a busca tendo em vista os hábitos do usuário na rede. Assim, conceberam arquitetura da informação como o desenho da organização, etiquetado, navegação e sistemas de busca para ajudar as pessoas a encontrarem e gerirem a informação mais adequadamente (LÓPEZ; GAGO; PEREIRA, 2003, p.198). Para os autores espanhóis López, Gago e Pereira, arquitetura da informação envolve a organização estrutural do mapa de conteúdo, os espaços internos que sustentam o sistema e o aspecto externo de um produto digital ou de um *cibermedio*, como eles preferem denominar (*ibid.*).

<sup>41</sup> A liberação do pólo de emissão altera o modelo um-todos que caracteriza os meios tradicionais de comunicação de massa, instaurando um modelo todos-todos, ou qualquer um para qualquer um. Sobre esta característica propiciada pelas novas tecnologias de comunicação e informação, ver Pierre Lévy (1993, 1999), André Lemos (1997; 2003) e, especificamente no campo do jornalismo, ver Barrie Gunter (2003).

concepção de narrativa, também em função da utilização, cada vez mais freqüente, das bases de dados para organizar, armazenar e estruturar os conteúdos.

### **1.2.1 Uma estética: modos diferenciados para visualizar a informação**

A introdução de computadores, das redes de computadores e da web criou um novo paradigma para as práticas de catalogar e arquivar o conhecimento. A vasta quantidade de informação digitalizada na cultura contemporânea fez com que novos modos para se acessar e recuperar a informação fossem considerados, uma vez que os sistemas organizacionais estabelecidos por bibliotecas e museus já não se mostravam adequados.

Situam-se precisamente na segunda metade da década de 90 as teorizações acerca do reconhecimento, por parte dos artistas, do poder conceitual e estético das bases de dados. É quando elas são percebidas como um suporte adequado para se empreender o esforço artístico, pois configuravam-se como um território rico para os artistas trabalharem nos mais variados níveis. “Bases de dados e arquivos servem como comentários prontos sobre nosso viver social e político contemporâneo”<sup>42</sup>, registrou a teórica e artista Victoria Vesna (1999, 2000a, p. 171).

À semelhança de experiências seminais que intentaram criar novos sistemas para armazenamento, compartilhamento e recuperação da informação, como aquelas realizadas por Vannevar Bush com o Memex (Memory Extension), na década de 40, ou

---

<sup>42</sup> “Databases and archives serve as ready-made commentaries on our contemporary social and political lives”. (t.A.).

o projeto Xanadu, de Thed Nelson, em 1965, a Alexandria Digital Library (ADL) e o Corbis Image Library, nos anos 90, os artistas, destacou Vesna, tinham uma grande oportunidade na conjuntura histórica para desempenhar importante papel na definição e no desenho de sistemas de acesso e recuperação da informação. Mais ainda: podiam empregar as tecnologias digitais para criar, constituindo novas práticas e objetos artísticos, com os quais responderiam, à sua maneira, aos modos de representar a organização do conhecimento da vida contemporânea.

Ademais, os artistas que naquela altura trabalhavam com mídia digital, particularmente na rede, estavam já conscientes de que o excesso de informação e o *design* de navegação estavam se tornando elementos de uma prática estética:

Indeed the development of context in the age of information overload is the art of the day. This is particularly true of the current artistic practice on the net in which artists frequently coopt and summon work and data of others. One of the by products of a "global culture" is the emergence of meta-structures, which include physical architectures, software such as the browser technology that allows us to view information on the Internet via the WWW, and artworks that are meta-art pieces, including work of not only other artists but the audience itself. Artists working with the net are essentially concerned with the creation of a new type of aesthetic that involves not only a visual representation, but invisible aspects of organisation, retrieval, and navigation as well. Data is the raw form that is shaped and used to build architectures of knowledge exchange and as an active commentary on the environment it depends on-the vast, intricate network with its many faces (VESNA, 1999, 2000a, p. 173-174)<sup>43</sup>.

---

<sup>43</sup> “Certamente o desenvolvimento do contexto na era sobrecarregada de informação é a arte do dia. Esta é a verdade particular da prática artística atual na rede, na qual artistas freqüentemente cooptam e convocam trabalho e dados dos outros. Um dos subprodutos de uma “cultura global” é a emergência de metaestruturas, as quais incluem arquitetura física, *software* tal como a tecnologia *browser* que nos permite ver a informação na internet via a WWW e trabalhos artísticos que são pedaços de metaarte, incluindo trabalhos não apenas de artistas, mas da própria audiência. Artistas que trabalham com a net são essencialmente concernentes com a criação de um novo tipo de estética que não envolve somente uma representação visual, mas também aspectos invisíveis de organização, recuperação e navegação. O dado é a forma crua/bruta que é modelada e usada para construir arquiteturas de trocas de conhecimento e como um comentário ativo sobre o vasto ambiente do qual ele depende, rede intrincada com suas variadas faces”. (t.A.).

Prosperavam, assim, as idéias em torno de uma estética específica da sociedade computadorizada, ou da cibercultura: a estética base de dados. Artigos versando sobre as práticas, assim como os aspectos desta estética foram reunidos na coletânea *Database Aesthetics: Issues of Organization and Category in Online Art*, publicada na *AI & Society: The Journal of Human-Centred and Machine Intelligence*, de 1999, a qual Victoria Vesna integrou e editou ao lado de Karamjit S. Gill, David Smith e Janet Vaux. O artigo dela, intitulado *Database Aesthetics: Of Containers, Chronofiles, Time Capsules, Xanadu, Alexandria and the World Brain*, teve uma versão posterior, *Database are us*, que saiu no volume 14 da *AI-Society* de Maio de 2000, juntamente com os demais trabalhos da edição de 1999, entre eles o de Lev Manovich, *Database as a Genre of New Media*, e o de Eduard Kac, *Time Capsule*.

Naquele mesmo ano, Steve Dietz (2000, 2004) começava a sua investigação sobre o uso de bases de dados na criação artística e, na versão de número 4 do seu *Database Imaginary: Memory\_Archive*, assinalava que um arquivo em base de dados podia ser entendido como um equivalente para a “museificação”, embora alguns artistas fossem céticos quanto a essa possibilidade, em função da natureza vibrante da *net culture* que eles estavam criando. Contudo, salientava que o esquema classificatório inerente a qualquer base de dados estava para a arte assim como a ornitologia está para os pássaros e, por isso, havia muitos artistas adotando as BDs como o seu meio, como uma estética e plataforma para desenvolver seus conceitos. Não apenas tomando-as como meros recipientes de metadados sobre eles (DIETZ, 2004, p.01).

Para a investigadora e artista multimídia Lucia Leão (2003), a lógica das bases de dados conduz a ansiarmos por uma saída, por uma ousadia, por uma estética subversora. Neste sentido, defende que o desenvolvimento de um trabalho em novas

mídias deve ser compreendido a partir da criação e organização de um banco de dados, bem como a elaboração de uma interface, sendo esta criada como um mapa, uma chave de acesso para as informações do sistema (LEÃO, 2003, p.89).

Já Christiane Paul (2004) se refere à expressão estética base de dados como um termo que tem se tornado um *slogan* do domínio digital e que constitui questões semânticas interessantes sobre o campo da pesquisa e da arte relacionadas com bases de dados. Mas o que exatamente quer dizer estética base de dados? A resposta, Paul nos oferece no artigo *The Database as System and Cultural Form: Anatomies of Cultural Narratives* (2004). Ela explica que, no discurso sobre arte digital, o termo é freqüentemente usado para descrever os princípios estéticos aplicados na imponente lógica das bases de dados para qualquer tipo de informação, filtrando coleções de dados, e visualizando os dados (PAUL, 2004, p.01).

Neste sentido, afirma, estética base de dados geralmente se torna uma forma conceitual e cultural potencial, um modo de revelar caminhos visuais de conhecimento, crenças e comportamento social. Os dados em si, no sistema que contém seus recipientes, não guardam qualquer valor estético. Pelo contrário, deles poder-se-ia dizer que estão mais para uma antiestética, pois são agrupados em tabelas e estruturas, que, a despeito de carregarem significado limitado, têm, porém, o potencial para múltiplas conexões relacionais. Ainda sobre o termo estética base de dados, Christiane Paul esclarece:

The term is seldom used to refer to the aesthetics of the database as structure itself, although it certainly implies that meaning and the structure of a database is inherently connected to the results produced by the filtering of the data contained in it and the nature of its visualization (PAUL, 2004, p.01)<sup>44</sup>.

---

<sup>44</sup> “O termo é raramente usado para referir à estética das bases de dados como estrutura, apesar de ele certamente implicar que o significado e a estrutura de uma base de dados estejam inerentemente conectados com os resultados produzidos pela filtragem do dado contido nela e na natureza da sua visualização”. (t.A.).

Os modelos de dados sempre contêm uma camada de algoritmos<sup>45</sup> matemáticos e de conceitos. Significa dizer: qualquer arte digital ou objeto da nova mídia – que se processa, cria, produz por meio do computador – tem uma camada de código e de algoritmos<sup>46</sup>. Esses produtos são construídos por instruções de *softwares*, empregados para a criação e a manipulação.

Os produtos que circulam no meio digital sempre consistem de um ‘*back end*’, formado pelos algoritmos e conjuntos de dados que permanecem escondidos, e de um ‘*front end*’, que é a parte visível experienciada pelos usuários. Os resultados gerados podem alcançar desde complexos visuais até processos abstratos de comunicação, pois, enquanto a arte digital é predominantemente visual, outros trabalhos focam mais sobre o dado bruto e o modelo base de dados em si. O entendimento comum da estética base de dados, explica Christiane Paul (2004, p.03), é mais relacionado com as operações do ‘*front end*’ – abrangendo as manifestações visuais e implicações culturais do conceito de algoritmo – do que com as do ‘*back end*’, dos dados e sua estrutura (PAUL, 2004, p.03).

Para Steve Dietz (2000, 2004), as bases de dados se tornam mesmo uma espécie de “*back end*” da contemporaneidade, uma interface cultural ubíqua. “Even when there

---

<sup>45</sup> Na matemática e nas ciências da computação, um algoritmo significa um pequeno procedimento para resolver um problema recorrente. Ele representa os passos necessários para se realizar uma tarefa, e a sua implementação pode ser feita por um computador ou por um ser humano. Diferentes algoritmos podem realizar a mesma tarefa usando um conjunto diferenciado de instruções em mais ou menos tempo, espaço ou esforço do que outros. Por exemplo, um algoritmo para se vestir pode especificar que você vista primeiro as meias e os sapatos antes de vestir a calça, enquanto outro algoritmo especifica que você deve primeiro vestir a calça e depois as meias e os sapatos. Neste caso, o primeiro algoritmo é mais difícil de executar que o segundo. Um programa de computador pode ser considerado como um algoritmo explícito. (WHATIS e WIKIPEDIA).

<sup>46</sup> De acordo com David Berlinsky (2000), o algoritmo é a segunda grande idéia científica do Ocidente, levada totalmente à consciência apenas na segunda metade do século XX, principalmente com o trabalho de Alan M. Turing. “Foi ele quem introduziu a idéia de máquina imaginária, um dispositivo que transformava várias abstrações em uma única construção brilhantemente simples” (ibid., p.230). Turing foi influenciado pelas formulações elaboradas ainda no século XVIII por Kurt Gödel, Gottfried Leibniz, Alonzo Church e Emil Post. Berlinsky define algoritmo como um procedimento eficaz, um modo de fazer uma coisa em um número finito de passos discretos. “Um conjunto de regras, uma receita, uma prescrição para a ação, um guia, uma diretiva concatenada e controlada, uma intimação, um código, um esforço feito para jogar um complexo xale verbal sobre o caos inarticulado da vida” (ibid., p.16).

does not exist, technically speaking, a collection of data, or information, that is specially organized for rapid search and retrieval, the potential of getting-retrieving-finding what you want is omnipresent, just on the other side of the interface” (DIETZ, 2000, 2004)<sup>47</sup>.

Em geral, a estética de uma base de dados é intrinsecamente relacional, tanto do ponto de vista dos dados em seus recipientes, como do relacionamento estabelecido pelo componente do *software*. Indica, pois, possibilidades diversas para se traçarem os processos individuais, culturais e comunicativos. Ademais, cabe notar que, como um princípio subjacente à estrutura de qualquer objeto da nova mídia, como refere Paul (2004), as BDs delineiam um campo aberto que inclui desde uma rede como a internet (uma gigantesca base de dados), até um simples e particular conjunto de dados.

Artistas em diferentes países têm apresentado criações assentadas na estética base de dados. Seja como *web sites*, instalações, videoinstalações, entre outros projetos que podem ser representados em formatos variados. O elo comum entre eles é o emprego de possibilidades inerentes à mídia digital para a classificação e – em alguns casos a reconfiguração de materiais existentes – em uma estrutura de base de dados.

O exemplo pioneiro é o projeto *File Room*<sup>48</sup>, de Antonio Muntades (1994-2004), que mostra como as instâncias de censura têm ocorrido ao redor do mundo e passa a acontecer de forma diferenciada, segundo um modelo bidirecional de fluxo de informação, filtragem colaborativa, múltiplos pontos de vista e transgressão de fronteiras geográficas.

---

<sup>47</sup> “Mesmo quando não existe, tecnicamente falando, uma coleção de dados, ou informação, que seja especialmente organizada para rápida pesquisa e recuperação, o potencial para obter-recuperar-encontrar o que você quer é onipresente, do outro lado da interface”. (t.A.).

<sup>48</sup> <http://www.thefileroom.org/>

O efeito da rede, escreveu Steve Dietz (2000, 2004), permite a autoria distribuída de casos de censura e cria, no caso de *File Room*, a possibilidade de um relacionamento assimétrico entre poder e controle.



Figura 1 - *Projeto File Room*, de Antonio Muntadas, 1994-2004: casos de censura ao redor do mundo. Out.2006.

Outro projeto bastante referenciado é o *Time Capsule*<sup>49</sup>, de Eduardo Kac, apresentado em São Paulo<sup>50</sup>, em 1997, como um evento-instalação. Kac demonstrava todo procedimento de implantação de um *microchip* com um número de identificação programado integrado a uma espécie de bobina e capacitor no seu corpo. Esse

<sup>49</sup> [http://vv.arts.ucla.edu/publications/publications/98-99/ai\\_society/kac.html](http://vv.arts.ucla.edu/publications/publications/98-99/ai_society/kac.html)

<sup>50</sup> São Paulo tem sediado muitas exposições e eventos sobre arte e tecnologia. O Instituto Itaú Cultural (<http://www.itaucultural.org.br/>) é um dos mais atuantes neste aspecto. Em setembro de 2006, o instituto realizou a terceira bienal de arte e tecnologia, com o evento *emoção artificial 3.0 interface cibernética* (<http://www.itaucultural.org.br/>), com mostra de trabalhos artísticos e simpósios que contou com a participação, entre outros, dos referenciados Edmond Kouchot e Golan Levin, além de brasileiros como Diana Domingues, cujo trabalho integra arte, informática e automação industrial. Júlio Plaza, Giselle Beighelman, Monica Tavares, entre outros, são nomes do cenário nacional envolvidos com a produção de poéticas digitais.

implante era então escaneado e gerava um sinal de rádio, que energizava o *microchip* e transmitia um código numérico. Como parte do procedimento, Kac registrava a si mesmo numa base de dados. O objetivo do trabalho era demonstrar que os humanos poderiam se tornar extensões do computador, ou de uma interface base de dados, pois os seus corpos passam a ser assunto para classificação e, atualmente, controlam várias bases de dados que estão além de nosso alcance corpóreo.

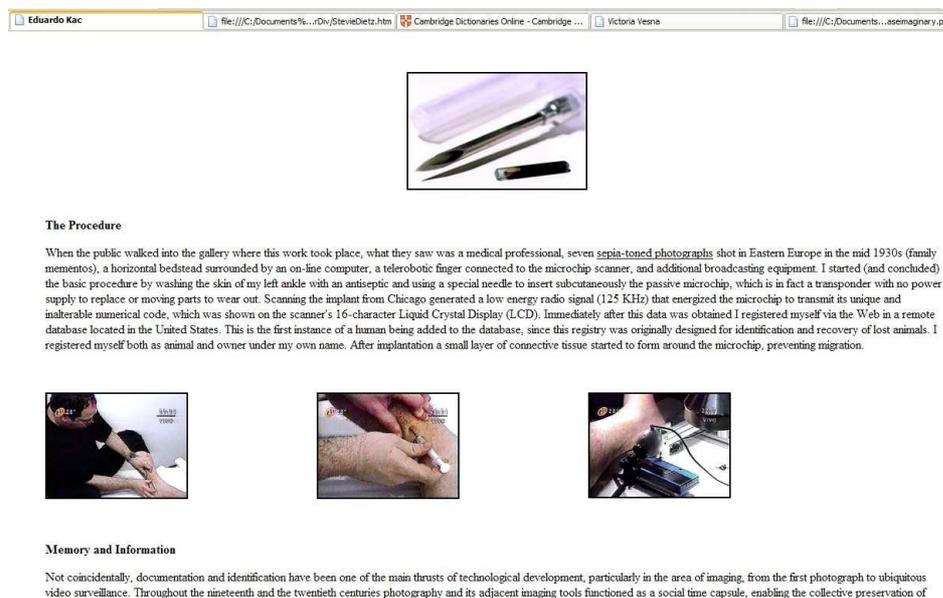


Figura 2 -*Time capsule*, evento-instalação de Eduardo Kac: corpo como interface base de dados. Out.2006.

Victoria Vesna, em parceria com programadores, *designers* e cientistas, tem produzido instalações, projetos web, entre outros, nos quais são realizadas representações do “corpo de dados” – *data body* – respondendo ao chamamento que ela própria fazia no artigo *Database are us*: “Em uma era na qual nós estamos crescentemente conscientes de nós mesmos como base de dados, identificados por números de segurança social e estruturas genéticas, é imperativo que artistas participem

ativamente em como o dado é modelado, organizado e disseminado” (VESNA, 1999, 2000a)<sup>51</sup>.

Um dos trabalhos é *Datamining Bodies in Ruhr*<sup>52</sup>, de 2000, no qual o usuário/participante é convidado a extrair – em 333 segundos – um corpo de dados a partir das sete camadas propostas para o *Ruhr Body*, no qual ele vai descender de fluidos fabris alimentados com dados continuamente atualizados. O dado, neste caso, resulta da mistura entre imagens de carvão extraído da era industrial, de imagens médicas do corpo e de sons, conforme a explicação conceitual de Vesna sobre o projeto.

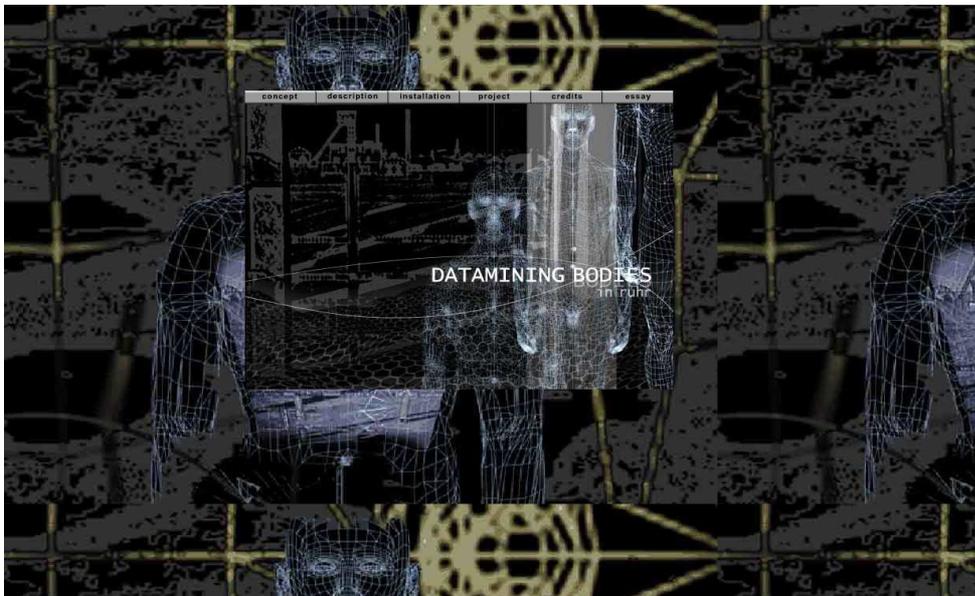


Figura 3 - Instalação e projeto web, *Datamining Bodies in Ruhr*, idealizado por Victoria Vesna. Out.2006.

<sup>51</sup> “In an age in which we are increasingly aware of ourselves as databases, identified by social security numbers and genetic structures, it is imperative that artists actively participate in how data is shaped, organised and disseminated”. (t.A).

<sup>52</sup> <http://notime.arts.ucla.edu/mining/mine.html>

Já *The Secret Life of Numbers*<sup>53</sup>, de Golan Levin, com Martin Wattenberg, Jonathan Feinberg, Shelly Wyncoop, David Elashoff, e David Becker, revela uma abordagem interessante sobre o que resulta do movimento entre o *back* e o *front end*. Eles se basearam em estudos sobre a popularidade contida entre os números de zero a um milhão e, usando um *software*, dados provenientes de mecanismos de busca, além de técnicas estatísticas, determinaram a popularidade dos números, expondo a sua vida secreta: caminhos de associações que refletem interesses culturais. A visualização consiste de duas interfaces, que combinam um histograma com a popularidade dos números e um gráfico no qual os conjuntos de números mais populares são exibidos. Um menu vai indicar o valor de popularidade para cada número escolhido, revelando as associações, por exemplo, com datas históricas, entre outros eventos quotidianos, desde 1997.

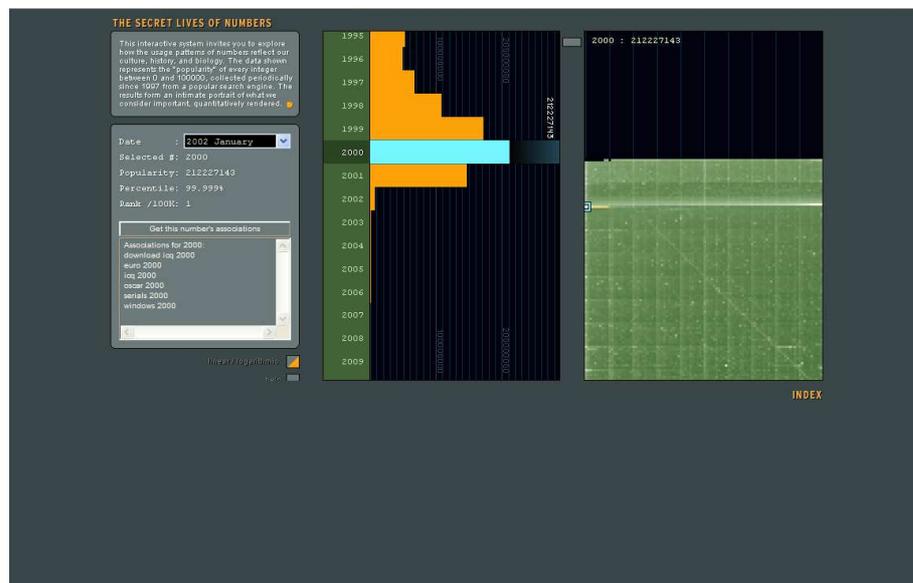


Figura 4 - *The Secret Life of Numbers*: popularidade dos números associados a vários eventos, desde 1997. Out.2006.

<sup>53</sup> <http://www.turbulence.org/Works/nums/applet.html>

A estética base de dados também é mote para exposições, como a 17<sup>a</sup> ArtFutura<sup>54</sup>, realizada em outubro de 2006, em dez cidades da Espanha, que elegeu a estética base de dados para mostrar as representações de um mundo construído pelo fluxo de dados<sup>55</sup>, e aquela considerada como emblemática, a *Database Imaginary*<sup>56</sup>. Esta última circulou por galerias do Canadá, Estados Unidos e algumas capitais europeias, entre 2004 e 2006.

Naquela exposição, os curadores (entre eles, Steve Dietz) incluíram 23 trabalhos realizados por 33 artistas, entre 1971 e 2004. Para eles, o período expandido contemplaria trabalhos que vinham sendo feitos desde a década de 70, pois foi quando o termo base de dados começou a ser usado, a computação começou a se expandir e passou-se a ter acesso amplo a vastas quantidades de informações organizadas em BDs. Entre os trabalhos, estavam: *Visitors' Profile*, de Hans Haacke (1971), *Databank of the Everyday*, de Natalie Bookchin (1996), *The File Room*, de Antonio Muntades (1994-2004), *Data Mining the Amazon*, de Andie Waller (2002), *Soft Cinema: Mission to Earth*, de Lev Manovich (2004), e *Zapped!*, do grupo Preemptive Media (2005).

Aqui, cabe comentário adicional sobre o trabalho apresentado por Manovich, pois integra um projeto maior, o *Soft Cinema*. Trata-se de um trabalho de investigação sobre um paradigma diferenciado para a realização de um filme ou audiovisual. Em vez do *script*, propõe começar com uma ampla base de dados de centenas de videoclipes, imagens, sons, texto, cores e, então, gerar narrativas a partir deles. A intenção é reimaginar o cinema como algoritmo, o qual tem procedimentos para governar como as

---

<sup>54</sup> <http://www.artfutura.org/>.

<sup>55</sup> A exposição foi destaque na edição do dia 28/10/2006 no *Elmundo.es*.  
<http://www.elmundo.es/navegante/2006/10/27/tecnologia/1161936287.html?a=6a18a6aa166fe2881ed58b67897c617c&t=1162059832>.

<sup>56</sup> <http://databaseimaginary.banff.org/index.php>

cenas contidas nas bases de dados serão re combinadas, daí resultando visualizações diferenciadas de imagens correspondentes às várias cenas e, em vez de um único final, a abordagem guiada pelo software permite apresentar várias versões diferentes. O DVD *Soft Cinema: Navigating the Database*, lançado em 2005, em parceria com o designer Andreas Kratky, traz, além de *Mission to Earth*, os filmes *Absences* e *Texas*.



Figura 5 – *Soft Cinema*. Uma das cenas de *Mission to Earth*: alegoria sobre a imigração na Guerra Fria. Out.2006.

Outro exemplo que podemos citar é o *The Dumpster*<sup>57</sup>, criado por Golan Levin, Kamal Nigam e Jonathan Feinberg (2005-2006). O objetivo do projeto de visualização interativa *on-line* é retratar a vida romântica dos adolescentes norte-americanos, intento realizado através da filtragem dos dados de 20 mil *posts* reais sobre os finais dos relacionamentos, obtidos através de inúmeros *blogs*. Eles coletaram frases indicativas de final de namoro e, a partir dos cruzamentos dos dados obtidos, revelaram os motivos, o sexo e a idade do autor, seu estado emocional, entre outros aspectos. *The Dumpster* está afinado com o que se chama “*social data browser*” (MANOVICH, 2006), pois permite

<sup>57</sup> <http://www.tate.org.uk/netart/bvs/thedumpster.htm>

navegar sobre detalhes íntimos das experiências das pessoas e de grupos sociais, tendo-se o particular e o geral apresentados simultaneamente.

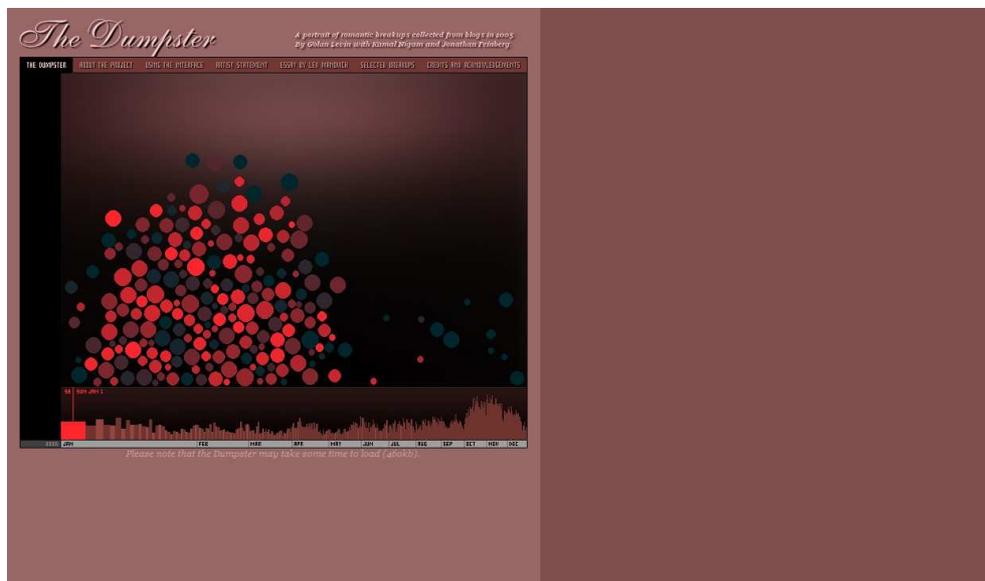


Figura 6 - *The Dumpster*: estratégias de análise de dados mostram novos tipos de representação social. Out.2006.

A estética base de dados, seus princípios plásticos e modos particulares de revelar caminhos visuais de conhecimento possuem um potencial singular para aplicação na área específica do jornalismo digital, pela originalidade e nível de diferenciação que trazem para a apresentação dos conteúdos. Alguns produtos e projetos que circulam na web já estão se apropriando de técnicas que permitem novas possibilidades para se trabalhar a visualização de informações jornalísticas.

Entre os casos referenciais estão: o *Newsmap*<sup>58</sup>, uma aplicação que reflete visualmente as atualizações do *Google News*<sup>59</sup>, site agregador de notícias; o

<sup>58</sup> <http://www.marumushi.com/apps/newsmap/newsmap.cfm>

<sup>59</sup> <http://news.google.com/?hl=en>

*Newzbubble*<sup>60</sup>, que rastreia informações de diversos *web sites* jornalísticos mostrando as notícias mais populares na forma de ‘bolhinhas’ para cada assunto, tema ou personalidade em destaque; o *News Brief* e o *News Explorer*, ambos do *European Media Monitor*.

Todos eles operam automaticamente com algoritmos e aplicações para fazer a varredura de informações jornalísticas na WWW. No caso dos dois últimos exemplos, as informações são apresentadas na forma de resumos (*News Brief*), com *links* para se ler a íntegra dos textos nos *sites* originais de onde foram capturados e também oferecendo contextualização para os conteúdos tratados (*News Explorer*). O diferencial está na visualização: pode-se ver desde a intensidade – baixo, normal, alto – do acesso às *breaking news* (últimas notícias), saber qual o país aparece em maior destaque na maioria das notícias, as personalidades mais citadas nas notícias referentes a um determinado país, até descobrir a rede de relações que se estabelece em torno de uma determinada personalidade.



Figura 7 - *Newsmap*: demonstração visual de notícias geradas pelo *Google News*. Out.2006.

<sup>60</sup> <http://www.newzbubble.com/index.php>



Figura 8 – Newzbubble. Visualização de informações jornalísticas através de interface de ‘bolhinhas’. Out.2006.

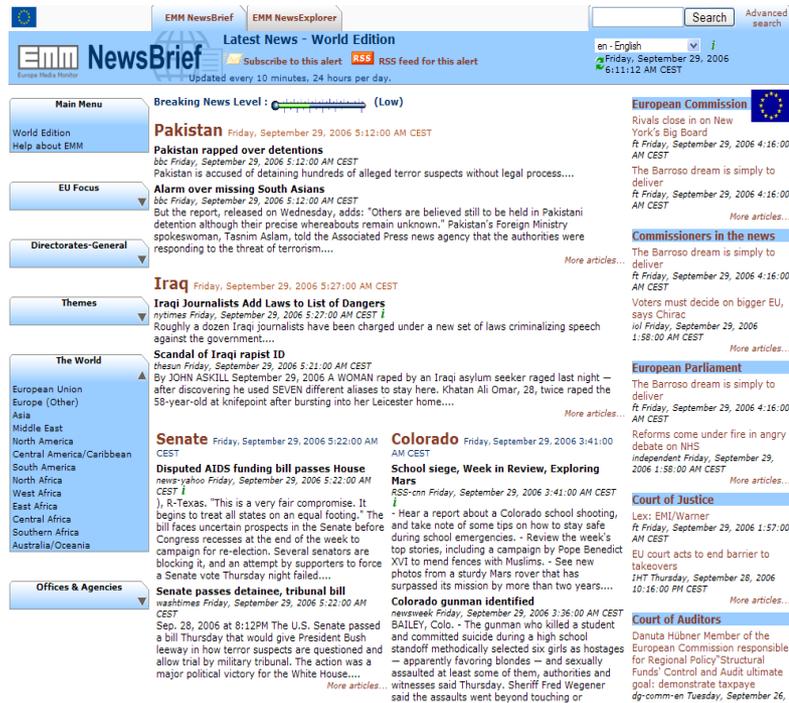


Figura 9 - News Brief do European Media Monitor. Out.2006.

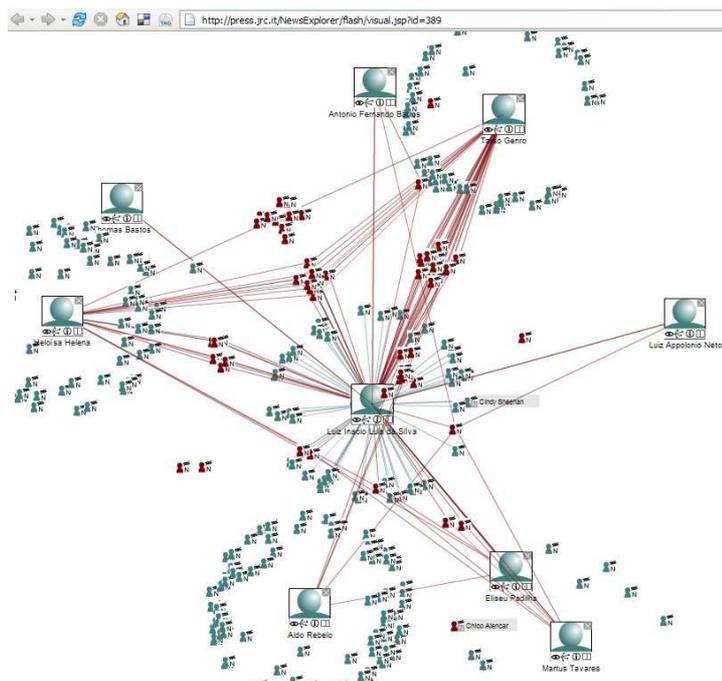


Figura 10 - No *News Explorer*, por exemplo, visualiza-se a rede de relações em torno do presidente Lula. Out.2006.

O blog brasileiro *Deu no jornal* também vem experimentando com um aplicativo para gerar visualizações para as relações estabelecidas entre as personagens e os assuntos das notícias publicadas. Há ainda outros exemplos de maneiras diferenciadas de se apresentar o conteúdo jornalístico aparecendo em seções ou canais específicos de um produto. É o que vem ocorrendo em *sites* de grandes conglomerados de comunicação, como a *BBC News*. A seção *Most Popular Now* permite não apenas saber as dez notícias mais populares num determinado momento do dia, mas também aquelas matérias e vídeos mais acessados em todo o *site*, quais foram destaques no dia anterior e como está o tráfego de acesso. Para isso, a interface oferecida é um mapa com os cinco continentes para que o usuário escolha a região do globo sobre a qual lhe interessa obter informações.

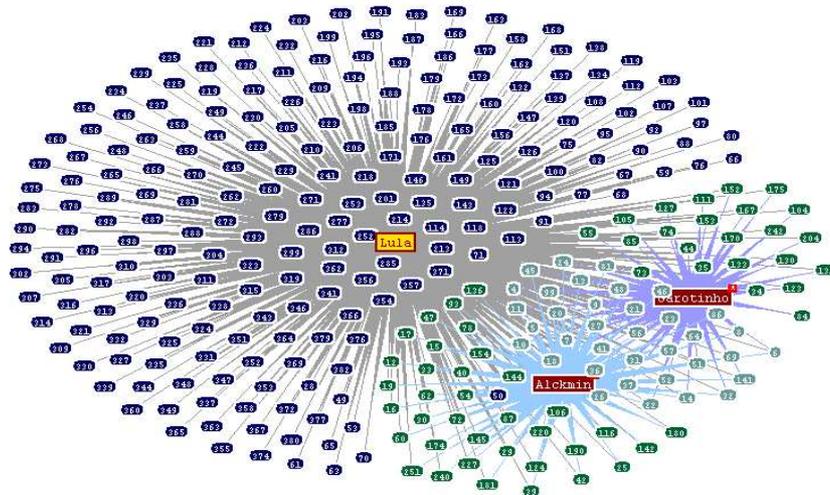


Figura 11 – No mapa do *Deu no Jornal*, se deve entrar com os nomes das pessoas ou assuntos sobre os quais se quer visualizar as relações. Dez.2006.

bbc.co.uk Home TV Radio Talk Where I Live A-Z Index Search

UK version International version About the versions Low graphics | Accessibility help

**BBC NEWS** OPEN BBC News in video and audio News services Your news when you want it

News Front Page World UK England Northern Ireland Scotland Wales Business Politics Health Education Science/Nature Technology Entertainment Have Your Say Magazine In Pictures Country Profiles Special Reports Programmes

**BBC News Most Popular Now**  
Friday, 29 September 2006

Most popular stories now Most popular around the site Most popular, day so far

**Worldwide, most popular stories now** Last updated 8 seconds ago  
This page updates automatically

MOST POPULAR	MOST E-MAILED
1 Canada 'sorry' over deportation	
2 US school gunman 'abused' girls	
3 What the US knows about visitors	
4 Itsy Bitsy writer 'death' error	
5 US Senate backs terror trial bill	
6 Death ends the myth of Tokyo Rose	
7 All you need is ubuntu	
8 Leak highlights a complex relation...	
9 Mistakes made in Iraq, says Straw	
10 Pakistan rapped over detentions	

Click on the map for a regional view

Site usage: Light, Normal, Heavy

Return to most popular stories worldwide

**TRAFFIC MONITOR** Worldwide traffic to the site is 1% above normal.

Low graphics/accessible version About this information

PRODUCTS AND SERVICES: E-mail news, Mobiles, Alerts, News feeds, Interactive TV, Podcasts

© BBC MMVI Most Popular Now | The most e-mailed story right now is: Beijing's penis emporium Back to top ^^^

Figura 12 - Seção *Most Popular* da *BBC News* usa técnicas diferenciadas para apresentar informações. Out.2006.

Nos capítulos 3, 4 e 5 examinaremos o aspecto da visualização, entre outros relacionados às funcionalidades e às categorias das bases de dados aplicadas ao jornalismo digital, com as devidas exemplificações.

### 1.2.2 Base de dados como ferramenta de comunicação

É através de técnicas como o *data mining* ou mineração de dados<sup>61</sup> que as bases de dados são transformadas numa poderosa ferramenta de comunicação da sociedade contemporânea, de acordo com a argumentação formulada por Luciane Lucas em sua tese de doutorado (LUCAS, 2004). Técnicas como esta permitem que as informações extraídas das bases de dados se transformem em conhecimento, devido à sua capacidade de revelar tendências, padrões e relações ocultas entre os dados. Com isso, os sistemas adquirem maior capacidade de planejamento e decisão, o que é essencial nas aplicações comerciais, especialmente naquelas do chamado *business intelligence*.

Como mostramos neste trabalho, não é só o mundo corporativo que tira partido das técnicas de *data mining*. Também na arte, na comunicação, potencialmente no jornalismo (COLLE, 2002, 2005a, 2005b; LIMA JÚNIOR, 2004; 2005; 2006; FIDALGO, 2005) e, inclusive, nas esferas do poder, o emprego do *data mining* vai se

---

<sup>61</sup> *Data mining* é uma forma automática de aquisição de conhecimento, uma técnica que torna um sistema inteligente à medida que auxilia no processo de decisão sem a interferência humana. Trata-se de um processo não-trivial de identificar, em dados, padrões válidos, potencialmente úteis. Sérgio Navega (2002) explica que esse processo se vale de algoritmos que processam os dados para encontrar os padrões válidos, novos e valiosos. Porém, ele ressalta que o *data mining* ainda requer uma interação muito forte com os analistas humanos, pois, em última instância, são eles os principais responsáveis pela determinação do valor dos padrões encontrados. No âmbito da inteligência artificial, conhecimento é definido como aquilo que torna um sistema capaz de decidir. Nisso estão inseridos o *data mining*, as redes neurais, a lógica *fuzzy* (introduzir incerteza nos sistemas inteligentes como maneira de forçá-los a conviver com a dúvida, assim como os humanos), a computação evolutiva e os agentes inteligentes.

consolidando como instrumento fundamental para a prospecção de conhecimento embutido nos dados. Essa é a parte do *data mining* conhecida como KDD (*Knowledge Discovery in Databases*) e sobre a qual Luciane Lucas (2004) dará maior ênfase, embora também aborde o eixo da modelagem de prognóstico, que permite predizer comportamentos de um grupo de consumidores, como a troca de uma marca por outra (*churning*).

Ao tomar as bases de dados como objeto de estudo e as técnicas de *data mining* como recorte em seu trabalho doutoral defendido na ECO/UFRJ, Lucas (2004, p.18) procura demonstrar que a tecnologia se tornou o principal alicerce para as decisões comunicacionais na sociedade contemporânea. É a partir de tendências e padrões identificados em bases de dados que campanhas de fidelização, segmentação de mercado e personalização de produtos e mensagens são realizadas. Estratégias, campanhas e instrumentos são desenhados a partir de informações que ela classifica como preditivas (prognósticas) e que vão definir desde o melhor *target* para responder a uma mala direta até a proximidade desejável de produtos na gôndola de um supermercado.

As possibilidades de extração de conhecimento em um *database* são infinitas: evolução dos perfis de clientes, tendências no comportamento de consumo, probabilidade de *churning*, padrões de fraude, predição de flutuações nas vendas de produtos sazonais, entre outros fatores que podem influenciar na seleção de mídias, na intensidade da comunicação ou na natureza da mensagem (LUCAS, 2004, p.18).

A disseminação das BDs como ponto de partida para estratégias de relacionamento com o consumidor é um fator que, segundo Luciane Lucas (2004, 2006), evidencia a transformação das práticas discursivas, uma vez que atualmente se experimenta um novo uso da informação. Lucas vai se apropriar das formulações de

Michel Foucault, entre outros, para reconhecer as bases de dados, os sistemas inteligentes e as ferramentas de extração de conhecimento como novos dispositivos de poder. São, segundo diz, instrumentos próprios de uma sociedade de controle<sup>62</sup>.

No caso específico do *data mining*, trata-se de uma força de comunicação do invisível, pois o padrão de discurso deste dispositivo é justamente ‘fazer ver e falar’ a informação que está oculta – informação que, em si, não existe, mas é produto da combinação de pares de relações. A mudança de paradigma no campo da comunicação é identificada por ela por meio de duas evidências teóricas: 1) é o invisível, o não-dito, que dá ordem e sentido ao visível, e 2) mais do que a informação prospectada, o que vale é o capital de nexos (ou seja, relações, grifo nosso) que uma informação irradia ao longo de si.

Advém daí a hipótese de Luciane Lucas (2004, 2006) de considerar o *data mining* como um dispositivo de algoritmização do cotidiano para proceder a “personalização em massa”. Essa algoritmização do cotidiano, conforme explica, torna-se cada vez menos ficcional à medida que os dados correntes e diários dos indivíduos são submetidos ao garimpo de *softwares* inteligentes.

Vivemos uma crescente e silenciosa algoritmização do cotidiano cuja lógica se instala na sociedade de consumo. E enquanto discutimos a força dos tradicionais veículos de comunicação ou mesmo o alcance das novas mídias, uma lacuna importante tem se formado: a do entendimento da lógica que ampara as práticas discursivas contemporâneas, no momento exato em que novas formas de saber, de conhecimento da alteridade, redefinem os contornos das práticas de poder (LUCAS, 2006, p.111).

---

<sup>62</sup> Nesta perspectiva, ela se aproxima da abordagem empreendida por Mark Poster (1995, 2000), que explora a relação das BDs com a questão cultural da formação do sujeito a partir da noção foucaultiana de discurso. Ou seja, no modo como as bases de dados atuam para a constituição de sujeitos fora da imediatez da consciência. Para Poster, as BDs operam como um *super-panopticon*, pois, como a prisão, as bases de dados operam de forma contínua, sistemática e sub-reptícia, acumulando informação acerca dos indivíduos e compondo-os em perfis (POSTER, 1995, p.69).

De maneira mais abrangente, ela considera pertinente falar de uma algoritmização do conhecimento, que passa a ser, nesse contexto, compressão do conhecimento construído ‘no’ e ‘sobre’ o mundo. Trata-se, pois, de uma prática discursiva diferenciada que reconfigura o próprio sentido de discurso. Ou, dito de outro modo: a algoritmização do conhecimento se desenha, no cenário contemporâneo, como uma prática discursiva cuja forma de enunciação é a própria compressão do discurso.

Através das proposições de Luciane Lucas (2004, 2006), buscou-se identificar o papel das bases de dados como ferramenta de comunicação, percebendo o nível de transformação que opera nas estratégias de comunicação para o relacionamento com os consumidores e, claro, algumas das conseqüências que acarreta. Mas, principalmente, pretendeu-se também apontar como as técnicas de *data mining* são empregadas para a descoberta de conhecimento através de dados ocultos, o que, para o jornalismo digital, pode trazer vantagens para a produção de conteúdos mais contextualizados e diferenciados.

### **1.2.3 Uma hipernarrativa: a narrativa interativa**

Se na vigência da forma cultural base de dados, uma estética BD é identificada, também encontramos outra concepção sendo redefinida – ou alargada – por conta de tal tecnologia: o conceito de narrativa. Para explicá-lo melhor, recorreremos, uma vez mais, às formulações de Lev Manovich (2000, 2001, 2005), por ser quem propõe a redefinição do conceito, bem como às repercussões delas nas proposições aportadas em trabalhos de outros autores.

Na definição de Manovich, narrativa ou hipernarrativa (numa analogia ao hipertexto) é aquela que resulta da soma de múltiplas trajetórias efetuadas através das bases de dados. O usuário da hipernarrativa, explica ele, atravessa as bases de dados seguindo os *links* entre os registros, conforme estabelecido pelo criador da base de dados. Logo, a narrativa tradicional linear é uma entre muitas outras trajetórias possíveis, constituindo uma escolha particular dentro de uma hipernarrativa (MANOVICH, 2001, p.227).

Explicando melhor, na cultura do computador, a narrativa é construída pela linkagem entre os elementos de uma base de dados em uma ordem particular, que define uma trajetória conduzindo elementos de um ponto para outro. Num nível material, sustenta o autor, a narrativa é apenas um conjunto de *links*, pois os elementos em si permanecem armazenados em uma base de dados (MANOVICH, 2001, p.231). Para demonstrar a pertinência da sua formulação, ele aplica a teoria semiológica que prevê as dimensões sintagmática (relações se dão de maneira explícita) e paradigmática (relações ocorrem de maneira implícita) para compreender os relacionamentos entre os elementos de um sistema.

Se no âmbito da linguagem, entre outros, os relacionamentos se dão segundo aquele esquema, na nova mídia esse relacionamento se inverte. Ou seja, segundo Lev Manovich (2001, p.231), a base de dados que até então cumpria um papel implícito, adquire existência material (paradigma), enquanto a narrativa, de explícita, passa à condição de desmaterialidade (sintagma). O paradigma, portanto, é privilegiado sobre o sintagma. A narrativa se torna virtualizada e a base de dados passa a existir materialmente, ou seja, se torna real.

Num exemplo de uma tela contendo alguns ícones, clicar sobre cada ícone conduz o usuário para uma tela diferente. No nível de uma tela individual, as escolhas

do usuário formam um paradigma para ele. No nível do objeto, o usuário se torna consciente de que está seguindo uma trajetória possível entre muitas outras. O que significa dizer que o usuário está selecionando uma trajetória de um paradigma de todas as trajetórias que são determinadas.

Interactive interfaces foreground the paradigmatic dimension and often make explicit paradigmatic sets. Yet they are still organized along the syntagmatic dimension. Although the user is making choices at each new screen, the end result is a linear sequence of screens that she follows. This is the classical syntagmatic experience (...) a new media user creates a sequence of screens by clicking on this or that icon at each screen (...) (MANOVICH, 2001, p.232)<sup>63</sup>.

É esta hipernarrativa que rege os modos como os produtos da nova mídia são explorados, colocando o usuário como um performer de ações para levar a narrativa adiante. Pois, se ele não se movimenta, seja clicando sobre os *links* e escolhendo o que lerá ou verá em seguida, ou não faz qualquer movimento para um *game* prosseguir, a narrativa pára<sup>64</sup>.

Conforme o autor russo, naturalizado americano e professor da Universidade da Califórnia (San Diego/EUA), a melhor maneira de se pensar nas relações entre narrativa e base de dados é compreendê-las como dois impulsos criativos básicos, duas respostas essenciais para se entender o mundo. Pois tanto uma como outra produzem inúmeros objetos híbridos, sendo mesmo difícil não encontrar traços de narrativa e de bases de dados num mesmo produto. Por outro lado, enciclopédias multimídia, museus virtuais,

---

<sup>63</sup> “Interfaces interativas colocam a dimensão paradigmática em primeiro plano e geralmente tornam explícitos os conjuntos paradigmáticos. Eles são ainda organizados junto à dimensão sintagmática. Embora a usuária esteja fazendo escolhas a cada nova tela, o resultado final é uma seqüência linear de telas que ela segue. Esta é a experiência sintagmática clássica (...) um usuário da nova mídia cria uma seqüência de telas ao clicar sobre este ou aquele ícone em cada tela (...)”. (MANOVICH, 2001, p.232).

<sup>64</sup> Sobre a questão da criação de fluxos narrativos multilíneares a partir de escolhas dos usuários, veja-se Palacios (1999).

CD-ROMs de artistas, bases de dados de bibliotecas, indexadores web e, claro, a própria web são demonstrações de que as bases de dados são mais populares atualmente do que foram antes.

Ao acolher as formulações de Manovich (2001) para avaliar a relação entre narrativa e base de dados em alguns produtos da nova mídia, Christiane Paul (2004) vai evidenciar a existência de componentes da narrativa tradicional (que requer o protagonista e o narrador, o texto e a história, e os relacionamentos de causa e efeito para eventos aparentemente desordenados) sobressaindo-se sobre uma estrutura base de dados.

Segundo verifica, tais componentes organizam os elementos constituintes que compõem hipertextos ou *games* (os quais são abertos e interativos) e, por isso, ocorre a ênfase da narrativa sobre a estrutura base de dados, que pode suportar uma narrativa, mas não se presta ela própria para gerar uma. Contudo, assegura Paul (2004), as bases de dados conferem a elas próprias uma categorização da informação e narrativas que podem ser filtradas para criar metanarrativas sobre a construção e detalhes culturais específicos do material original (PAUL, 2004, p.07).

Apart from the visualizations of data sets in general, which often do not explicitly reveal the database as an underlying structure, a multitude of projects has been tracing cultural subtexts by means of arranging information in a database that then become a means of revealing meta-narratives about cultural values and conventions (PAUL, 2004, p.07)<sup>65</sup>.

---

<sup>65</sup> “À parte as visualizações de conjuntos de dados em geral, os quais geralmente não revelam explicitamente a base de dados como uma estrutura subjacente, uma multitude de projetos tem traçado subtextos culturais por meio de informação organizada em uma base de dados a qual se torna um modo revelador de metanarrativas sobre valores e convenções culturais”. (t.A.).

A argumentação de Paul pode ser aplicada, por exemplo, no caso do projeto *The Dumpster*, que, a partir das informações contidas em bases de dados, criou caminhos para mostrar aspectos correspondentes do comportamento adolescente no momento de rompimento de relacionamentos amorosos. Ou, em consonância com Christiane Paul, metanarrativas que revelam traços comuns, entre outros dados, referentes a um grupo social, indicando valores e algumas convenções.

A ressalva que a autora faz é que bases de dados e narrativas não são formas mutuamente exclusivas, e, nesse ponto, reforça aquilo já afirmado por Manovich acima: objetos da nova mídia são produzidos a partir de bases de dados e, neles, traços da narrativa tradicional estão presentes, resultando em muitos híbridos. O aspecto diferencial a apontar é que os modos como se exploram tais objetos produzem um tipo novo de narrativa – a hipernarrativa – e, por outro lado, também geram as metanarrativas, que são criadas a partir da filtragem das informações organizadas nas bases de dados. Paul (2004) também vai notar que há, em muitos dos projetos de arte, uma imposição da lógica da base de dados sobre as formas da narrativa tradicional, pois tal lógica guarda maior afinidade com a raiz da mídia digital.

Elias Machado (2004a, 2004b, 2006) e Beatriz Ribas (2005) também empregaram a concepção de narrativa interativa ou hipernarrativa desenvolvida por Lev Manovich (2001) para pensá-la em relação à composição de narrativas jornalísticas no ciberespaço<sup>66</sup>. Nas investigações de ambos está presente a defesa de uma redefinição para o conceito de narrativa, que, no caso do jornalismo digital, significa a incorporação das bases de dados como formato e do espaço navegável como suporte para a criação de narrativas interativas.

---

<sup>66</sup> O termo foi cunhado por William Gibson no livro *Neuromancer* (1984), considerado marco inaugural do gênero cyberpunk, o primeiro movimento cultural a refletir criticamente sobre as implicações trazidas pelas novas tecnologias da informática.

Para Machado (2004b, 2006), a narrativa se configura cada vez mais como uma viagem através do espaço constituído pelos conjuntos estruturados de itens organizados na forma de banco de dados e se torna um conjunto contínuo de ações narrativas e explorações. Ele argumenta que, ao contrário da narrativa moderna, na qual o ouvinte, o leitor ou o telespectador acompanha a narração (ouvindo, lendo, vendo) sem interferir na lógica interna das ações, o fluxo da narrativa no ciberespaço, mais do que incorporar, vai depender diretamente da intervenção do usuário, ou do teleator.

A ação deste teleator, quando acessa, por exemplo, o espaço navegável de uma publicação jornalística, será a de intervir na narrativa. “Ao eleger como território de exploração um dos muitos módulos disponíveis e optar por uma, entre as várias linearidades propostas, ele desenvolve uma ação que interfere no curso da narrativa, que deixa de ser único como na narrativa jornalística convencional” (MACHADO, 2004b, p.04).

Beatriz Ribas (2005, p.06), por sua vez, propõe como definição da narrativa na web, a estrutura que incorpora os elementos da narrativa tradicional em associação aos elementos, características e princípios do meio digital, para desempenhar as funções de: 1) organizar e tornar facilmente acessíveis os dados na tela, tornando-se, desta maneira, a mediadora entre computador, usuário e produtor (JOHNSON, 2001; SCOLARI, 2004); e 2) criar ambientes diferenciados para as relações entre os dados, permitindo experiências distintas e possibilitando a identificação de diferentes tipos de produtos e de estratégias comunicacionais.

Os elementos básicos para os sistemas interativos são constituídos pela lexia, a interface e o banco de dados. Ribas (2005, p.04) os detalha, explicando que a lexia<sup>67</sup> é a unidade mínima de informação, ou seja, da narrativa, podendo conter o texto, a imagem,

---

<sup>67</sup> Na teoria do hipertexto desenvolvida por George P. Landow (1997), a lexia é a unidade mínima da narrativa, ao passo que, na narratologia, o texto é a unidade mínima da narrativa.

o som, ou ser uma combinação de todos; a interface é que faz a mediação, ou seja, permite a interação entre o usuário e o computador e varia de acordo com as diferentes configurações dos dados armazenados por uma BD; e o banco de dados reúne registros, ou eventos, que podem ser recuperados e relacionados de diversas maneiras, simples ou complexa, seqüencial ou fragmentada, linear ou multilinear, para serem acessados pelo usuário.

As características são aquelas definidas como específicas para o jornalismo digital: hipertextualidade, interatividade, multimídia, personalização, memória/arquivo e atualização contínua (PALACIOS, 2002a; 2003b; 2003c; 2004; MACHADO, 2003; ZAMORRA, 2002; MIELNICZUK, 2003; DIAZ NOCI & SALAVERRÍA, 2003; SALAVERRÍA, 2005b), as quais serão descritas no capítulo 3, e os princípios são aqueles pensados por Manovich (2001) para diferenciar a nova mídia, apresentados no subtópico 1.1.2.

Beatriz Ribas (2005, p.05) afirma haver uma relação de semelhança entre os elementos da narrativa tradicional e os da narrativa na web, explicitados segundo o esquema abaixo:

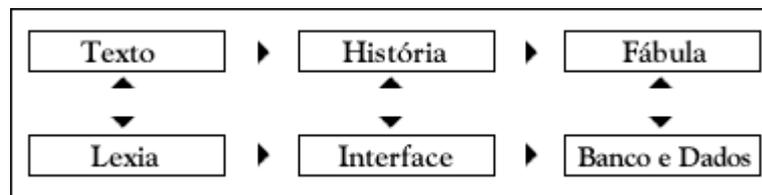


Figura 13 - Relação de semelhança entre os elementos. Fonte: Beatriz Ribas, 2005.

São as características de hipertextualidade, interatividade, multimídia, personalização, memória/arquivo e atualização contínua, que vão determinar como funcionarão as relações entre os elementos da narrativa tradicional e os elementos da

narrativa na web, ou narrativa hipertextual. Neste sentido, cabe notar: o ambiente interativo hipertextual demanda certos tipos de relações entre produtor, interface e usuário com os dados, que mudam a cada acesso ou a cada reconfiguração da interface (RIBAS, 2005, p.06).

O emprego de tais características na composição das narrativas jornalísticas – ou das peças informativas –, ao lado da gradual incorporação das bases de dados como formato pelas empresas de comunicação ou *sites* noticiosos de modo geral, concorrerá para tornar os conteúdos mais diferenciados e afinados com a narrativa no ciberespaço. E, por se tratar de uma narrativa interativa, como afirma Lev Manovich (e com o qual estamos de acordo), ela é constituída pelos *links* entre os registros de uma base de dados, permitindo, assim, que mais de uma trajetória seja possibilitada ao usuário.

O *link*, por sua vez, é o recurso que organiza as publicações digitais, permite o inter-relacionamento entre os conteúdos e estrutura as narrativas, assim como ativa as relações estabelecidas com o conteúdo por parte do usuário, deste com a publicação e com os demais atores, potencializando o funcionamento do hipertexto (MIELNICZUK; PALACIOS, 2001; MIELNICZUK; MARQUES, 2006).

A este usuário, ou teleator, como prefere denominar Elias Machado (2004b, 2006), cabe, ainda, o papel de manter a qualidade da narrativa interativa, seja como: 1) um ativador do fluxo que institui o espaço navegável, ou 2) como ator que interfere no desenvolvimento da narrativa propriamente dita e nas relações com os demais atores. Embora alguns possam temer tal atuação do usuário, por acarretar o desaparecimento do autor, reiteramos, com Machado, que a narrativa interativa torna a função do autor mais difícil do que antes da interação digital devido à necessidade de reagir às modificações decorrentes das ações dos teleatores e das reações dos personagens (2004b, p.05).

Sobre os modos de interação do usuário com os conteúdos, há certos princípios e estágios (MEADOWS, 2003) para guiá-lo no desenvolvimento da interação e ajudar na avaliação deste desenvolvimento, assim como também há, por outro lado, níveis diferenciados para que eles atuem como geradores de informação, no contexto da produção de conteúdos no jornalismo (CABRERA, 2005). O quadro a seguir traz as especificações dos princípios e estágios elencados por Mark Meadows (2003, p.39-44) e os níveis de participação definidos por Maria Ángeles Cabrera (2005, p.332).

MARK S. MEADOWS (2003)		MARIA A. CABRERA (2005)
03 PRINCÍPIOS	04 ESTÁGIOS	PRODUTOR DE INFORMAÇÃO 04 Níveis
<p><b>Ingresso/Saída</b> Todo ingresso ou saída de informação no sistema deve gerar mais informação;</p> <p><b>Dentro/Fora</b> Estabelece que deve existir diálogo entre os mundos interno e externo;</p> <p><b>Aberto/Fechado</b> Quanto mais o sistema for usado, melhor será seu funcionamento. Sistemas abertos, como a internet, são mais complexos e menos previsíveis.</p>	<p><b>Observação</b> Acomodação do usuário no ambiente. Ele se acostuma com o entorno da informação e com a maneira como ela está organizada (visão macroscópica);</p> <p><b>Exploração</b> Usuário cria mapa cognitivo que o leva a explorar os caminhos disponíveis (visão microscópica);</p> <p><b>Modificação</b> Envolve o processo de resposta/<i>feedback</i>. Usuário percebe que sua ação provoca algum tipo de mudança no sistema, pois obtém resposta às ações que realiza. Leva ao último estágio;</p> <p><b>Mudança recíproca</b> Usuário interfere no ambiente, que se modifica. A partir da resposta obtida, usuário volta a interferir no sistema. Processo cíclico.</p>	<p>Participação do usuário na produção dos conteúdos está relacionada com a atividade, a emissão, a personalização, a profundidade, a liberdade. Aspecto aberto e interativo da rede permite maior protagonismo por parte dos usuários. Essa participação se dá em distintos níveis de implicação:</p> <p><b>Protagonista</b> Implicação máxima do usuário por ser o protagonista do fato noticioso;</p> <p><b>Testemunha</b> Quando atua como observador direto;</p> <p><b>Colaborador</b> Aporta dados de contexto sem necessidade de ser testemunha direta;</p> <p><b>Comentarista</b> Opina sobre o conteúdo da informação.</p>

Quadro 2 - Princípios, estágios e níveis de participação do usuário

A inserção dos usuários nos processos de elaboração dos conteúdos jornalísticos, em especial no jornalismo digital, é defendida por muitos autores (MACHADO, 2002, 2003, 2004a, 2004b, 2006; PALACIOS, 2002a, 2003b, 2003c; BOWMAN, WILLIS, 2003; GILLMOR, 2004, 2005; BOCZKOWSKI, 2004; ALVES, 2005) e está colocada

como um ponto-chave do jornalismo participativo (ROSEN, 2003; 2005a; 2005b; 2006a; 2006b). Nos capítulos subseqüentes, voltaremos a discutir o assunto.

Com a abordagem realizada neste primeiro capítulo, pretendeu-se identificar o “lugar” das bases de dados na sociedade contemporânea, apresentando aspectos inerentes a esta tecnologia da informação e alguns dos modos como ela se insere e opera em diferentes áreas. Uma vez percorrido este caminho no capítulo que abre esta tese, prosseguiremos com as discussões mais específicas sobre as bases de dados no jornalismo e procedendo ao mapeamento das especificidades que nos levam a considerá-las como o aspecto-chave do jornalismo digital na etapa de transição entre a terceira e a quarta gerações de evolução, com a conformação do que denominamos **Jornalismo Digital em Base de Dados (JDBD)**.



## **AS BASES DE DADOS E O JORNALISMO**

## 2

**As bases de dados e o jornalismo**

---

*Perhaps the best way to describe the potential of this electronic information system to journalism is to suggest that online information databases will bring the benefits of the eighteenth-century print revolution to late twenty-first-century journalism.*

- Tom Koch.

*I feel we are in a database era in journalism, but it may become so common that we will no longer call it that in the future.*

- Bruce Garrison.

*Electronic databases mean that we have to add a knowledge of 'search' software to the set of skills that we've always applied to archives, record-offices and libraries. It also means that there are new ways of telling stories and making them relevant to our readers on an individual basis.*

- Jim Hall.

## 2.1 HISTÓRICO DE ADOÇÃO SEGUE CURVA DO “S”

Os computadores começaram a fazer parte da indústria do jornal no final da década de 60, quando se descobriu que, assim como os setores empresarial e comercial poderiam tirar partido da idéia de *time-sharing*, isto é, permitir a uma série de usuários explorar as capacidades de um mesmo computador ao mesmo tempo, também os jornais se beneficiariam com a introdução daquela máquina no seu processo de produção. As vantagens para o campo do jornalismo estariam não apenas nas técnicas do *time-*

*sharing*<sup>68</sup> para se realizar as tarefas em menor tempo e com maior precisão, mas, sobretudo, a introdução do computador levaria a uma reformulação de todas as funções dentro da empresa jornalística. O ano de 1971 (SHEDDEN, 1998) foi o marco maior, pois concretizou os esforços, imprimindo a mudança do processo mecânico para os sistemas baseados no computador para a fabricação de jornais.

Inicialmente, ele foi usado para melhorar e agilizar o processo de composição e, pouco a pouco, foi incorporando-se em todas as etapas de produção, por meio da implantação dos sistemas editoriais<sup>69</sup>, os quais passaram a desempenhar funções desde a chegada das notícias, a elaboração redacional, a composição dos textos e dos títulos até a paginação. Em seguida, o computador chegou às etapas de circulação e de distribuição e também alcançou os setores de fotografia, com o processo contínuo de digitalização das imagens (SMITH, 1980; CARITÁ, 1987; VIANNA, 1992; SHEDDEN, 1998; BALDESSAR, 2003).

A informatização da imprensa representa, segundo o estudo de Anthony Smith (1980), a terceira revolução na comunicação, isto é, nas técnicas humanas para o armazenamento, processamento e circulação da informação. Antes do computador, os protagonistas das revoluções prévias foram a escrita e a imprensa de Gutenberg (SMITH, 1980, p.03). Esse novo modelo se irradia a partir dos Estados Unidos e de lá para os jornais japoneses, europeus, sul-americanos e alcançam os periódicos de todo o mundo entre os anos 70 e o final dos 80. Cada um a seu tempo e conforme as condições específicas vai aderindo à informatização e, com ela, os jornais vão também expandir a

---

<sup>68</sup> À época, o mercado de *mainframes* que suportava o *time-sharing* e o processamento *on-line* era controlado pela *IBM*, com o seu sistema 370 (SHEDDEN, 1998).

<sup>69</sup> De acordo com Caritá (1987, p.294), sistemas editoriais são aqueles de teleelaboração que desempenham todas as funções fundamentais, desde a chegada das notícias até a paginação com produção do protótipo de página fechada, com uma eventual incisão direta da placa de impressão. Dentro de uma empresa editorial, eles seriam como o terminal nervoso de um sistema global de informações, constituído por computadores, bases de dados, satélites e redes de transmissão.

distribuição de suas edições para outros países, através de satélites, que permitem transformar os sinais em prova positiva, para serem impressas em outros locais<sup>70</sup>.

No Brasil, a *Folha de S. Paulo* foi o primeiro a introduzir terminais de computador em suas redações, projeto que vinha sendo elaborado desde 1968 e foi efetivado em 1983. Como a *Folha*, *O Globo*, *O Diário Catarinense*, *A Tribuna*, *Zero Hora*, *O Estado de S. Paulo* também se adaptaram à nova era do jornalismo, assim como revistas, a exemplo de *Veja*. Na Bahia, houve um certo atraso, e só na década de 90 o computador adentrou nas redações. Foi o jornal *Feira Hoje*, de Feira de Santana, aquele que primeiro se informatizou. Na verdade, ele já nasceu sob esse novo modelo. Na capital, Salvador, o *Bahia Hoje* foi o outro exemplo de um jornal criado já completamente informatizado. Em meados dos anos 90, os jornais *Correio da Bahia* e *A Tarde* se informatizaram e, posteriormente, a *Tribuna da Bahia*<sup>71</sup>.

Para os jornalistas em si, a chegada do computador nas redações significou uma grande mudança na realização das suas funções, causando mesmo um efeito revolucionário nos modos como passaram a desempenhar as tarefas diárias, da apuração à redação até a edição do material. Contudo, essa ‘revolução’ chegou atrelada à necessidade de reciclagem e atualização constantes para corresponder às pressões de se produzir cada vez melhor, num cenário mais competitivo, no qual o jornalista deveria responder com trabalho de qualidade para atender às exigências das empresas, que sempre cobram bons resultados em períodos de tempo cada vez mais curtos. Além disso, o desemprego e a extinção de algumas funções, como a de revisor, foram

---

<sup>70</sup> Sobre a ampliação do mercado para os jornais e o processo de distribuição das edições usando satélites, ver Morgaine (1980).

<sup>71</sup> Atualmente, só os três últimos jornais estão em operação em Salvador.

conseqüências negativas do impacto da informatização das redações<sup>72</sup>, que, por outro lado, também criou outros postos e funções para os jornalistas, assim como estabeleceu as bases para mudanças que viriam a seguir com a internet e a web.

Ao iniciar a abordagem neste segundo capítulo sobre o histórico de adoção das bases de dados no jornalismo, o fazemos com o breve relato da introdução dos computadores na indústria de jornais por serem os jornais os pioneiros no emprego desta tecnologia e sistema de informação, seguidos pelas revistas e emissoras de TV e de rádio. A informatização de todas as etapas do processo de produção dos jornais permitiu que, lentamente, *softwares* e outras ferramentas pudessem ser utilizados por jornalistas visando ao aprimoramento do seu trabalho diário de obtenção de notícias e à construção de relatos sobre os eventos quotidianos da sociedade para satisfazer as necessidades de informação dos cidadãos.

A introdução do computador no processo de produção é mesmo um passo crucial na evolução para a nova mídia e para o jornalismo digital, anos mais tarde (SHEDDEN, 1998). E as bases de dados tiveram um papel fundamental nessa história. Nora Paul (1999, p.109) as considera as responsáveis por terem trazido as primeiras doses da 'revolução da informação' para o jornalismo.

De acordo com alguns autores (GARRISON, 1998; HERBERT, 2000; REAVY, 2001; GUNTER, 2003), a informatização das redações e de todo o processo de produção jornalístico abriu caminho para a concretização daquilo que há muitos anos se esperava: trazer o poder dos computadores para o jornalismo para conferir maior qualidade aos métodos de apuração e à análise de informações. No entanto, foi preciso

---

<sup>72</sup> Sobre a informatização da imprensa, ver Morgaine, (1972), Smith (1980), e, especificamente sobre os impactos que ocasionou para os jornais e os jornalistas brasileiros, ver os trabalhos de Lins da Silva (1988), Vianna (1992) e Baldessar (2003).

esperar até a década de 70 para que tal possibilidade começasse a se concretizar, ainda que com os complicados sistemas *mainframes*, grandes e difíceis de operar.

O jornalista Philip Meyer é considerado um dos precursores no uso do computador. Isso ocorreu ainda em 1966, quando ele se encontrou com um IBM 7090 no *Harvard Computation Center* (MEYER, 1999). De início, as duas tentativas falharam aos seus propósitos, porém, em 1967, a partir dos estudos sobre o método de pesquisa das ciências sociais, ele voltou a usar o computador (com a ajuda de professores e de técnicos) para analisar os resultados da pesquisa de opinião pública na área metropolitana de Detroit sobre a causa dos motins urbanos que ocorreram no verão daquele ano.

Foi quando Meyer, então repórter do *Detroit Free Press*, introduziu no jornalismo técnicas das ciências sociais para realizar o seu trabalho, dando início ao que ele denominou *Precision Journalism*, uma metodologia para a coleta de informações utilizando o método científico e as técnicas de pesquisa das ciências sociais<sup>73</sup> aliadas ao uso do computador. Por aquela cobertura, Philip Meyer e o *Detroit* receberam o *Prêmio Pulitzer*, em 1968, o que levou Meyer a ser chamado pela *Newsweek* de “a computer reporter” e o tornou um dos mais solicitados para ensinar a outros jornalistas tais técnicas. O Jornalismo de Precisão foi a forma prévia da *Computer Assisted Reporting* (CAR) ou Reportagem Assistida por Computador<sup>74</sup>, que emergiria no início dos anos 90.

As bases de dados começaram a ser utilizadas nas redações a partir da década de 70, quando se descobriu o uso potencial desta tecnologia para o jornalismo. As primeiras utilizações como ferramentas para análise jornalística de arquivos oficiais

---

<sup>73</sup> Em 1973, Meyer lançou o livro-referência *Precision Journalism*, descrevendo as aplicações desta metodologia no jornalismo e, 18 anos depois, uma outra edição atualizou a obra: *The New Precision Journalism* (1991).

<sup>74</sup> Também chamada de *Computer Assisted Journalism* (CAJ) ou Jornalismo Assistido por Computador (JAC). Nesta tese, optamos pelo uso do termo CAR (em inglês) ou RAC (em português).

foram feitas, em 1972, pelos repórteres Donald L. Barlett e James B. Steele, do *Philadelphia Inquirer*, no estudo sobre o sistema judicial penal da *Philadelphia* (MEYER, 1993, 1999; SHEDDEN, 1998).

Contudo, o caso mais marcante foi o surgimento do *The New York Times Information Bank*, a primeira base de dados de um jornal (SMITH, 1980; COLOMBO, 1991; SHEDDEN, 1998; SQUIRRA, 1998). O *The New York Times* começou a gestar o projeto ainda em 1966<sup>75</sup> e, em 1973, o concretizou, sendo o pioneiro na percepção de quão valiosos poderiam ser os arquivos (considerados até então como peças mortas) para o seu negócio. Além do arquivo propriamente dito, ele continha resumos e uma variedade de citações originárias de outros jornais e publicações especializadas. O serviço foi criado com a intenção de atender, principalmente, os usuários externos, vendendo informações para as grandes corporações, especialmente, além, claro, de servir ao consumo interno, como observa Anthony Smith (1980, p.114).

Personnel inside the New York Times Information Bank and outsiders renting usage on a time basis have to have recourse to microfiche or microfilm for any material more detailed than the abstract... It is one thing to design a system that enables reporters to quickly check the ages and birth dates of Congressmen and another thing to help businessmen learn the details of the policies or market research activity of rival companies. But in the longer run the newspaper's electronic morgue could develop as a new kind of public service, interacting with the domestic user and playing an important role within the "information society" (SMITH, 1980, p.114)<sup>76</sup>.

---

<sup>75</sup> No mesmo ano, o jornal *The Washington Post* também começou a desenvolver projetos de uso de bases de dados. Naquele período, o termo empregado pelo jornal era *datbank*.

<sup>76</sup> "A equipe do *New York Times Information Bank* e os de fora que alugam sobre tempo de uso semanal têm que ter o recurso para a microficha ou microfilme para qualquer material mais detalhado que o resumo... Ele é algo que cria um sistema que permite aos repórteres checar as idades e as datas de nascimento dos congressistas e, entre outras coisas, ajuda os homens de negócio a aprenderem os detalhes das políticas ou pesquisa de atividade de mercado das companhias rivais. Mas, no longo caminho dos arquivos eletrônicos, os jornais poderiam desenvolver um novo tipo de serviço público, interagindo com o usuário doméstico e desempenhando um importante papel dentro da "sociedade da informação" (t.A.).

De modo generalizado, o entendimento inicial para as bases de dados é aquele que as vê como sinônimos de estocagem, armazenamento estruturado de informações nos computadores para fácil recuperação. Daí porque o seu emprego primordial tenha recaído sobre a organização dos arquivos e a conseqüente abertura destes para acesso público. Os usuários externos passaram a ter acesso a textos memorizados, artigos produzidos pelas redações dos jornais e por outras fontes. O acesso era realizado na própria sede dos jornais por meio de terminais de vídeo<sup>77</sup> e, a seguir, por meio da internet, que iria promover a descentralização do uso dos arquivos computadorizados.

No final da década de 70, as bases de dados já eram uma realidade nos maiores jornais dos Estados Unidos, chegando depois às redações das TVs. Índices de bases de dados, BDs textuais e bibliográficas voltadas para assuntos das ciências físicas, da medicina, da área jurídica eram os tipos de bases de dados disponíveis<sup>78</sup> para ser pesquisadas por meio de terminais ligados aos grandes computadores centrais, que funcionavam conectados à linha telefônica e a um *modem*.

---

<sup>77</sup> Eram denominados *Video Display Terminals (VDTs)*.

<sup>78</sup> Foi entre 1972 e 1973 que a informação começou a ser compilada em bases de dados através da iniciativa de empresas como *Dialog* e *LexisNexis* (GARRISON, 1998, p.72).



Figura 14 - Um dos terminais exibindo informações disponibilizadas pela empresa de BDs, *LexisNexis*. Fonte: [www.lexisnexis.com](http://www.lexisnexis.com). Set.2006.

Só num segundo momento vão ser mais amplamente incorporadas ao processo de apuração, como fonte para pesquisa e análise de dados estatísticos e a melhor contextualização das matérias e reportagens. Essa fase vai coincidir com um novo estágio, inaugurado pela computação pessoal, nos anos 80, quando surge o *Personal Computer (PC)* e, com ele, a microinformática, que demarca um período de novos desenvolvimentos e maior facilidade para se aprender a operar os computadores.

Também é naquele momento que os primeiros *softwares* de bases de dados são introduzidos nas redações, assim como o uso de planilhas eletrônicas. Os até então limitados terminais (chamados *dumb terminals*<sup>79</sup>) podiam fazer mais a partir dali: escrever histórias, checar bases de dados *on-line* governamentais e comerciais de empresas como *LexisNexis*, *Dialog* e *Dow Jones News Retrieval* – disponibilizadas

<sup>79</sup> “Terminais burros”, pois tinham poucos recursos, eram lentos e só suportavam texto. Esses computadores funcionavam por meio de um terminal de vídeo, conectado ao *mainframe*, usando linha telefônica e *modem*.

através dos provedores de acesso *CompuServe*, *Prodigy* e *America Online* – e também permitiam criar BDs próprias, analisar planilhas eletrônicas, entre outros programas (GARRISON, 1998, p.20-21). A década de 80 foi, enfim, de aprendizado e de incremento no uso das bases de dados pelos jornalistas (KOCH, 1991), principalmente para a elaboração de reportagens investigativas, muitas das quais premiadas com o *Pulitzer*<sup>80</sup>.

Anthony Smith (1980) e Tom Koch (1991) foram os primeiros a prever o potencial do uso das bases de dados como ferramentas no jornalismo. Examinando períodos distintos, mas relacionados entre si, ambos se referem aos bancos de dados eletrônicos *on-line* como os agentes de transformações fundamentais no processo de coleta, seleção e contextualização das notícias. Também verificaram que mudanças de atitude e habilidades seriam requeridas dos profissionais jornalistas para adequar posturas e posicionamentos diante de um novo recurso acrescido ao processo de produção da informação.

Enquanto Smith previu que o uso de várias BDs *on-line* pelas redações de veículos tradicionais mudaria o modo como os jornalistas coletavam, sintetizavam e recuperavam as notícias, Koch chegou mesmo a dizer que o melhor modo de descrever o potencial do sistema eletrônico de informação para o jornalismo seria considerar que as bases de dados de informação *on-line* trariam os benefícios da revolução da imprensa do século XVIII para o jornalismo do século XXI justamente pelo impacto e o nível de mudança que representariam. Segundo Tom Koch (1991, p.66), o sistema permitiria aos jornalistas de diários e semanários não apenas basear seus relatos sobre “o que foi dito

---

<sup>80</sup> Em 1987, por exemplo, como descreve Tom Koch (1991, p.278), o *Akron, Ohio, Beacon Journal* ganhou o *Prêmio Pulitzer* pela cobertura do caso da aquisição da *Goodyear Tire e Rubber Company* por um grupo financeiro europeu. Os jornalistas utilizaram vários recursos, incluindo bases de dados especializadas no setor de negócios e informações financeiras, como a *Disclosure*, para seguir os rumores e determinar o nível de certeza sobre os possíveis compradores, entre outros dados. O site do *Pulitzer* ([www.pulitzer.org](http://www.pulitzer.org)) traz a lista de todos os ganhadores do prêmio, desde 1917.

por X”, mas também colocar as declarações num contexto amplo a partir de um trabalho profundo de pesquisa prévio.

Uma outra novidade dos anos 80, relacionada ao emprego de bases de dados, é o videotexto – ou bases de dados videotexto –, lançado como a grande promessa de inovação e futuro para o jornalismo. O videotexto pode ser definido como o sistema de transmissão de informações por cabo ou sem fio que funciona por meio de um computador pessoal ou terminal especial, ou de aparelho de televisão conectado à rede telefônica por um *modem* ou aparelho decodificador de sinais analógicos em digitais e vice-versa.

Os experimentos para a entrega de notícias através da tecnologia do videotexto foram desenvolvidos durante os anos 70, com o lançamento comercial do serviço ocorrendo em 1979, a partir dos sistemas inventados na Europa, fundamentalmente o sistema britânico *Prestel* (RECODER, ABADAL, CODINA, 1995; GUNTER, 2003). O videotexto se expandiu durante a década de 80 e no começo da de 90 como um serviço de informação eletrônica, que, a despeito dos altos investimentos recebidos de grandes grupos e cadeias de jornais nos Estados Unidos (*Time*, *Time Mirror* e *Knight-Ridder*), entre outras empresas em países como Canadá, Japão, Alemanha, Bélgica e Espanha, teve maior desenvolvimento e aceitação na França, com o sistema *Minitel*<sup>81</sup>, que ainda é bastante utilizado pelos franceses.

---

<sup>81</sup> Além do videotexto, também havia o teletexto. Ambos eram sistemas de informações textuais transmitidas por cabo ou sem fios e enviadas para uma tela. Havia dois tipos de videotexto: o interativo (também chamado *viewdata*), transmitido através de redes telefônicas; e o de difusão (teletexto), transmitido por meio de ondas hertzianas. Os dois sistemas vigoraram em muitos países, como na Inglaterra, onde havia o videotexto *Ceefax* (BBC) e o teletexto *Oracle*. No Brasil, a RBS (*Rede Brasil Sul*, afiliada da *Rede Globo*) fez algumas experiências com o teletexto. Para saber mais sobre o sistema de videotexto e teletexto, ver os trabalhos de autores como Smith (1980); Recoder, Abadal, Codina (1995); Armañanzas, Díaz Noci, Meso (1996); Fidler (1997); Zamora (2002) e Gunter (2003). Em geral, esses sistemas são considerados os precursores do *web journal* ou *ciberperiódico*, como preferem denominar os autores espanhóis (DÍAZ NOCI, SALAVERRÍA, 2003; SALAVERRÍA, 2005a, SALAVERRÍA, 2005b; LÓPEZ, GAGO, PEREIRA, 2003; GARCÍA et al, 2005; LÓPEZ, 2006).



Figura 15 - O *Minitel I* foi um dos aparelhos utilizados para se acessar o sistema desenvolvido na França. Fonte: <http://en.wikipedia.org/wiki/Videotex>. Out.2006.

Enquanto isso, no Brasil, o videotexto foi introduzido em 1982 pela Telesp<sup>82</sup>, empresa de telefonia de São Paulo (SIQUEIRA, 1995, p.155), sendo seguida por grupos jornalísticos como o *Estado de S.Paulo*, além de outras empresas regionais de telefonia<sup>83</sup>. O sistema videotexto, porém, não alcançou o nível esperado de utilização, principalmente pelo limitado grau de interatividade, que seria satisfeito anos mais tarde com o surgimento da World Wide Web – a face hipertextual e gráfica da internet.

---

<sup>82</sup> Usando o Sistema Brasileiro de Videotexto (VDT) via Embratel, a então empresa estatal que controlava a telefonia no Brasil até 1995.

<sup>83</sup> Na Bahia, por exemplo, a Central de Videotexto da *Telebahia* alcançou certa popularidade no início da década de 90, oferecendo informações bancárias, auxílio à lista telefônica, informes em tramitação na Justiça, correio eletrônico e dados sobre os serviços prestados pela *Telebahia* a seus usuários, além do Vídeo Papo, o mais popular entre os serviços.



Figura 16 - Tela da Central de Videotexto do Sindicato dos Jornalistas de Santos/SP, a primeira central de videotexto fora da capital paulista. Fonte: <http://www.novomilenio.inf.br/ano97/97hista7.htm>. Out.2006.

Antes mesmo da web, na virada dos anos 80 para os 90, as bases de dados alcançaram novo período de expansão, emergindo como protagonistas da *Computer Assisted Reporting (CAR)*, definida por Bruce Garrison (1998, p.11) como a aplicação de computadores para obter informação para a apresentação das notícias, envolvendo o uso de computadores de todos os tamanhos, de *mainframes* a PCs até os computadores de mão. Àquela altura, os repórteres já utilizavam mais frequentemente os computadores e as ferramentas já tinham se tornado mais fáceis de ser manipuladas pelos jornalistas, que, assim, tinham menos resistência para adotá-los.

Mais do que uma variação original do termo Jornalismo de Precisão, o mais apropriado seria dizer, a partir da perspectiva de Nora Paul (1999, p.109), que foram as bases de dados que fundaram a Reportagem Assistida por Computador. Ambos os termos incluem dois níveis de utilização: 1) empregar os computadores para a busca de informação e recuperá-la de outros computadores e suas bases de dados e 2) usar os computadores para analisar bases de dados originais e bases de dados de outras fontes

de informação para os relatos jornalísticos. Alguns autores (SEMONCHE, 1993; GARRISON, 1998) chegaram a denominar esse modelo de *database journalism* ou jornalismo em base de dados.

A *CAR* teve duas fases de desenvolvimento: a primeira vai do começo até meados dos anos 90, e a outra começa na segunda metade da década, já marcada pelo uso da web entre outros recursos, como o *e-mail*, o sistema operacional *Windows95*, além de programas como o *PDF (Portable Document Format<sup>84</sup>)*, *PowerPoint*, planilhas eletrônicas, entre outros *softwares*.

A web – e a internet – desvela outro momento de expansão para as bases de dados, promovendo a sua descentralização e tornando-se ela própria a mais importante BD de informação nas redações. Entre 1997 e 1998, organizações públicas e privadas empreenderam grandes esforços para disponibilizar as suas respectivas bases de dados para pesquisa na WWW, enquanto várias empresas informativas passaram a incorporar ferramentas com BDs baseadas na web com acesso por meio das intranets (as redes internas das corporações). No final da década de 90, a Reportagem Assistida por Computador se popularizou, propagando o seu uso por jornais de todos os tamanhos, serviços de cabo, revistas e TVs.

Outro *turning point* para as bases de dados ocorreu também nos 90, quando surgiram as primeiras versões digitais dos jornais na web, inicialmente de modo experimental, em 1992 (SHEDDEN, 1998), oferecidas como serviços dos grandes provedores *America Online*, *CompuServe*, *Prodigy*. Entre os jornais, estavam o *Chicago Tribune*<sup>85</sup> e o *San Jose Mercury News*<sup>86</sup>, que, em seguida, lançaram seus respectivos *web*

---

<sup>84</sup> Formato de arquivo universal que permite distribuir documentos na web mantendo o seu formato original.

<sup>85</sup> <http://www.chicagotribune.com/>

<sup>86</sup> <http://www.mercurynews.com/mld/mercurynews/>

sites e estabeleceram as suas operações digitais desvinculadas dos provedores, assim como fizeram outros jornais, como *The New York Times*<sup>87</sup> e *The Washington Post*<sup>88</sup>.



Figura 17 - Home do *Chicago Tribune*, disponibilizado via *America Online*, em 1992.

Fonte: [http://iml.jou.ufl.edu/carlson/history/ScreenShots/chicago\\_online.jpg](http://iml.jou.ufl.edu/carlson/history/ScreenShots/chicago_online.jpg).

No Brasil, o *Jornal do Brasil*<sup>89</sup> foi o primeiro a disponibilizar a edição completa na web, em 1995, embora experiências anteriores tenham sido realizadas pelo *Estadão*<sup>90</sup> e pelo *Jornal do Commercio*<sup>91</sup>, de Pernambuco. Depois do *JB*, a *Folha de S.Paulo*<sup>92</sup>, O

<sup>87</sup> <http://www.nytimes.com/>

<sup>88</sup> <http://www.washingtonpost.com/>

<sup>89</sup> <http://www.jb.com.br>

<sup>90</sup> <http://www.estadao.com.br>

<sup>91</sup> <http://jc.uol.com.br/>

<sup>92</sup> <http://www1.folha.uol.com.br/fsp/>

*Globo*<sup>93</sup>, *Zero Hora*<sup>94</sup>, *Estado de Minas*<sup>95</sup> também lançaram seus *sites*. Na Bahia, *A Tarde*<sup>96</sup> e o *Correio da Bahia*<sup>97</sup> disponibilizaram as suas versões digitais em 1996.



Figura 18 - Em 1996, o *site* do *JB* já oferecia notícias em “tempo real”, com a criação da editoria “Extra”. Fonte: <http://jbonline.terra.com.br/destaques/2005/10anos/evolucao.html>.

A partir daquele momento, as bases de dados seriam agregadas aos *sites*, ampliando a oferta de informação, sobretudo com a disponibilização do material de arquivo, aliado ao recurso do hipertexto para a organização das informações. As bases de dados tornaram-se, portanto, uma vez mais, fundamentais para o jornalismo e para a

<sup>93</sup> <http://oglobo.globo.com/>

<sup>94</sup> <http://www.clicrbs.com.br/jornais/zerohora/jsp/default.jsp?uf=1&local=1&section=Home>

<sup>95</sup> <http://www.estaminas.com.br/>

<sup>96</sup> <http://www.atarde.com.br>

<sup>97</sup> <http://www.correiodabahia.com.br>

modalidade específica que surge com a internet e a WWW – o jornalismo digital. Elas chegaram ao ano 2000 sendo vistas não apenas como uma boa solução para armazenamento, recuperação e compartilhamento de informações, mas como fontes para a produção de matérias jornalísticas.

Nos primeiros anos deste século XXI, as BDs tiveram seu significado alargado, sendo consideradas como formato no jornalismo digital (MACHADO, 2004a, 2004b, 2006) e, gradualmente, adotadas pelas organizações jornalísticas como um pressuposto para adequar as suas estruturas à lógica do ciberespaço. Elas têm o potencial para garantir o diferencial desta modalidade jornalística em relação às anteriores (COLLE, 2002, 2005a, 2005b; FIDALGO, 2003, 2004), favorecer a emergência de um novo paradigma, bem como possuem o potencial para gerar uma nova metáfora (BARBOSA, 2004a; 2004b, 2004c, 2005a, 2005b, 2006a, 2006b). É neste contexto, portanto, que propomos o **Jornalismo Digital em Bases de Dados (JDBD)**.

Uma maneira para se compreender mais apropriadamente o histórico de adoção das bases de dados no jornalismo é por meio da teoria da difusão de inovações tecnológicas (ROGERS, 1995) e a sua curva do “S” – ou curva da adoção de inovação – tal como propõe Barrie Gunter (2003), e com o qual concordamos. A maneira como essa história vem sendo brevemente contada até aqui, por si, já nos remete à visualização de um cenário de utilização das BDs que delineia de fato um “S”. Além disso, cabe lembrar o que afirmou Tom Koch há 15 anos: “It is possible to argue that the diffusion of electronic database technologies is a certainty simply on the basis of the history of the diffusion of previous information-related innovations” (KOCH, 1991, p.301)<sup>98</sup>.

---

<sup>98</sup> “É possível sustentar que a difusão da tecnologia de bases de dados eletrônicas é simplesmente uma certeza sobre as bases da história da difusão de inovações prévias relacionadas à informação”. (t.A.).

Vejam, então, como a teoria da difusão de inovações pode ser apropriada para explicar o emprego das BDs no jornalismo.

### **2.1.1 Abordagem ancorada na teoria da difusão de inovações**

A teoria da difusão de inovações descreve como novas idéias ou práticas são introduzidas e adotadas em um sistema social, tendo um foco especial sobre as relações comunicacionais e fluxos de informação que promovem tal adoção. Esta teoria foi formalizada por Everett Rogers em 1962, no livro chamado *Diffusion of Innovations*, que, na sua quarta edição (ROGERS, 1995), revista e atualizada, recebeu este mesmo título numa tentativa de identificá-lo com a tradição seqüencial dos estudos de difusão. Difusão, como explica o autor, é o processo pelo qual uma inovação é comunicada através de certos canais entre os membros de um sistema social. Trata-se de um tipo especial de comunicação, no qual as mensagens são concernentes a novas idéias (ROGERS, 1995, p.05).

Inovação, ele esclarece, é uma idéia, prática ou objeto que é percebido como novo por um indivíduo ou outra unidade de adoção. Perceber a novidade de uma idéia implica uma reação por parte do indivíduo. Se a idéia parece nova para ele, ela é uma inovação. A novidade de uma inovação envolve novo conhecimento, mas pode ser também expressada em termos de persuasão ou uma decisão para adotar determinada inovação. Vale notar que as inovações a que se refere Everett Rogers estão relacionadas

diretamente com inovações tecnológicas, tanto que ele toma os dois termos (inovação e tecnologia) como sinônimos<sup>99</sup>.

Os quatro elementos principais da teoria da difusão de inovações são: a própria inovação, os canais de comunicação pelos quais se procede a sua divulgação, o tempo e o sistema social. Estes elementos são identificáveis em cada estudo da pesquisa de difusões e em cada campanha de difusão ou programa (ROGERS, 1995, p.11) e estão ilustrados na figura a seguir, que é a curva do modelo de adoção de inovação, mais conhecida como “S” *curve*. Ela é explicada da seguinte maneira: Difusão é o processo pelo qual (1) uma Inovação (2) é Comunicada através de Certos Canais (3) sobre o Tempo (4) entre os membros de um Sistema Social.

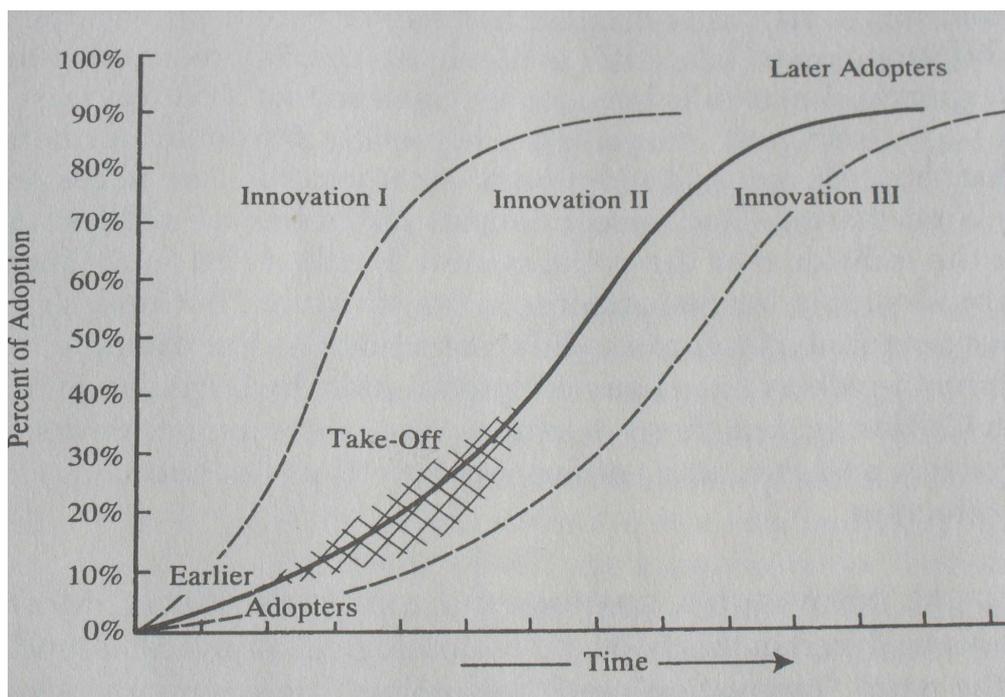


Figura 19 - Modelo da curva de adoção de inovação ou Curva do “S” explica o processo de difusão.  
Fonte: Rogers (1995)

<sup>99</sup> A referência, neste caso, é a definição de tecnologia como informação, formulada por Thompson (1967) e Eveland (1986), segundo descreve Everett Rogers: “Tecnologia é informação e transferência é um processo de comunicação e, assim, transferência de tecnologia é a comunicação da informação”. (ROGERS, 1995, p.12). Tecnologia geralmente compreende dois componentes, conforme Rogers assinala: o *hardware* (ferramenta que embute a tecnologia como um objeto material ou físico) e o *software* (consistindo a base da informação para a ferramenta).

Neste processo, há sempre aqueles considerados como *early adopters*, ou seja, os que aderem prontamente a uma certa tecnologia e exercem influência sobre outros membros do grupo, ou esses membros compartilham interesses similares com aqueles e então terminam por também adotar uma inovação. Há, ainda, os céticos quanto à aceitação de uma tecnologia, no entanto, em fases posteriores, segundo as pesquisas levadas a cabo dentro da teoria da difusão de inovações, eles passam a considerá-la e, então, se juntam àqueles que vão aderir um pouco mais tarde a uma determinada tecnologia.

Ondas sucessivas de adoção serão contínuas até a inovação alcançar um ponto máximo de expansão (ou saturação), formando massa crítica, que varia dependendo das características da inovação e do sistema social. O modelo da curva de adoção de inovação indica que, no início de sua utilização, uma tecnologia passa por uma fase de poucos atores aderindo a ela, e, num momento seguinte, ocorre uma aceleração, depois uma estabilização, podendo voltar a experimentar nova fase de propagação ou ter a sua expansão descontinuada.

Segundo a perspectiva de Everett Rogers (1995), difusão é um tipo particular de comunicação. Sendo assim, o autor define comunicação como um processo no qual os participantes criam e compartilham informação com os outros para alcançar um entendimento mútuo. Esta definição implica que comunicação é um processo de convergência (ou divergência), quando dois ou mais indivíduos trocam informação para mover-se em direção a uma outra (ou ficar à parte dela) pelo sentido que eles atribuem a certos eventos (ROGERS, 1995, p.17-18). No modelo convergente, olha-se para o que está anteriormente colocado num evento e o que se seguirá, percebendo que aquele evento é apenas uma parte de um processo total no qual a informação é trocada entre

dois indivíduos. Por exemplo: um cliente pode ir até o agente de mudança<sup>100</sup> com um problema e uma inovação é recomendada como uma possível solução. A interação entre o agente e o cliente pode continuar através de vários ciclos, como um processo de troca de informação.

A essência do processo de difusão, ele diz, é a informação trocada, por meio da qual o indivíduo comunica a nova idéia para um outro ou para vários outros indivíduos. Os *mass media*, em geral, foram assinalados por Rogers (1995, p.18) como os canais mais rápidos e eficientes para informar às pessoas a existência de uma inovação, tendo, por outro lado, afirmado que os canais interpessoais são mais efetivos para persuadir alguém a aceitar uma idéia nova. Também as redes de comunicação, como a internet, operam neste sentido, pois indivíduos interconectados seguem os caminhos dos fluxos de informação.

O terceiro elemento no processo de difusão – o tempo – é considerado pelo autor como um dos pontos fortes da sua teoria<sup>101</sup>. Essa dimensão temporal diz respeito ao processo de tomada de decisão de uma inovação, que envolve cinco passos: conhecimento, persuasão, decisão, implementação e confirmação; como e quando a inovação é adotada por uma pessoa ou unidade de adoção; e a taxa de adoção de uma inovação em um sistema. Essa taxa é medida pelo número de membros do sistema que adota a inovação em um determinado período de tempo. Já o sistema social – quarto elemento – é definido como um conjunto de unidades relacionadas que estão engajadas na solução de um problema para cumprir uma meta comum. Os membros ou unidades

---

<sup>100</sup> Um agente de mudança é um indivíduo, empresa privada, agência pública que influencia o cliente a aderir a uma inovação/tecnologia (ROGERS, 1995, p.27).

<sup>101</sup> Teoria esta que recebeu como uma das críticas a de ser tecnologicamente determinista por tratar a inovação como algo dado e estar focada mais sobre os efeitos ou o impacto das inovações em um sistema social (LIEVROUW, 2002). Por outro lado, a teoria de difusões foi e continua sendo muito empregada em diversas áreas e tem servido, algumas vezes, como *insight* para outras formulações como ocorreu com Roger Fidler em *Mediamorphosis* (1997), um processo denominado como tal para pensar a evolução tecnológica dos meios de comunicação.

de um sistema social podem ser indivíduos, grupos informais, organizações e/ou subsistemas, ou seja: comunidades, organizações, nações, mercado, indústria.

Everett Rogers (1995) assinala, ainda, que a propagação de uma inovação não deve ser analisada como um evento isolado, como fizeram alguns teóricos da difusão, pois mais apropriado é perceber que um conjunto de inovações é difundido ao mesmo tempo em um sistema social, e, portanto, deve-se tomá-las como interdependentes. Ademais, como ele mesmo afirma, a difusão de inovações é essencialmente um processo social no qual, subjetivamente, a informação percebida sobre uma nova idéia é comunicada. O significado de uma inovação é, deste modo, gradualmente elaborado através de um processo social de construção.

As cinco Características das Inovações é que vão ajudar a explicar como elas são percebidas pelos indivíduos e por que elas possuem diferentes taxas de adoção, assegura Rogers (1995, p.16-17). Essas características são: Vantagem relativa (*Relative advantage*), Compatibilidade (*Compatibility*), Complexidade (*Complexity*), Testabilidade (*Trialability*), Observabilidade (*Observability*). Na década de 70, a partir de resultados colhidos em pesquisas de campo, o investigador acrescentou uma outra característica àquelas que já integravam as suas formulações teóricas: a característica de Reinvenção ou *Re-invention*. As especificidades de cada uma delas estão descritas no quadro a seguir.

<b>Vantagem relativa</b> <i>(Relative advantage)</i>	<b>Compatibilidade</b> <i>(Compatibility)</i>	<b>Complexidade</b> <i>(Complexity)</i>	<b>Testabilidade</b> <i>(Triability)</i>	<b>Observabilidade</b> <i>(Observability)</i>	<b>Reinvenção</b> <i>(Re-invention)</i>
Nível em que uma inovação é percebida como melhor que a idéia que ela suplanta. Pode ser medido em termos econômicos, mas o prestígio social, a conveniência e a satisfação são também fatores importantes. Quanto maior é o nível de percepção sobre a vantagem de uma inovação, mais rápida será a taxa de adoção.	Nível em que uma inovação é percebida como consistente com os valores existentes, as experiências passadas, e as necessidades potenciais daqueles que irão adotá-la. Uma inovação que é incompatível com os valores e normas de um sistema social não será adotada tão rapidamente quanto uma que seja compatível.	Nível em que uma inovação é percebida como difícil para entender e usar. Algumas inovações são facilmente entendidas pela maioria dos membros de um sistema ou subsistema social, enquanto outras são mais complicadas e serão adotadas mais lentamente.	Nível em que uma inovação pode ser experimentada sobre uma base limitada. Novas idéias podem ser testadas sobre um plano de instalação e, geralmente, serão mais facilmente adotadas que inovações que não podem ser implementadas por partes. Uma tecnologia que pode ser testada representa menos incerteza para as pessoas que vão usá-la, pois se pode aprender testando-a.	Nível em que os resultados de uma inovação são visíveis para outros. A visibilidade estimula a discussão sobre uma inovação entre as pessoas e geralmente requer informação para avaliá-la, estimulando a sua adoção. Por exemplo, há inovações como o uso do computador nos lares, que são relativamente menos observáveis e, deste modo, o processo de adoção é mais lento.	Nível em que uma inovação é mudada ou modificada pelo usuário no processo de adoção e implementação. Pode ser medido pelo uso que uma pessoa faz de uma nova idéia a partir da versão original que foi promovida pelo agente de mudança. A flexibilidade de uma inovação/tecnologia favorece que os usuários possam implementá-la em modos diversos. Porém, esclarece Rogers, uma inovação não é necessariamente invariante no seu processo de difusão, e adotar uma inovação tal como recomendada não necessariamente significa ser passivo, apenas decide-se seguir aquele padrão ou <i>template</i> para aquela nova idéia.

Quadro 3 – Características das Inovações Fonte: Rogers (1995, p.16-17)

A teoria da difusão de inovações tecnológicas tem sido aplicada em diversos campos do saber, e, aqui, consideramos apropriado tomá-la como referencial para perceber o histórico de adoção das bases de dados no jornalismo, desde o momento inicial, localizado na década de 70.

Inferimos, a partir da perspectiva de Barrie Gunter (2003, p.92), que o emprego da tecnologia de bases de dados no jornalismo – especialmente na modalidade jornalística presente no ciberespaço – encontra-se num renovado movimento ascendente da curva do “S”, experimentando, assim, nova fase de expansão, devido ao seu potencial para a gestão de produtos jornalísticos digitais. Só que, agora, o uso das BDs é pensado e orientado para a estruturação, a organização e a apresentação das informações, observando-se um conjunto de funcionalidades, além das categorias que conferem especificidade e estabelecem as bases de dados como um paradigma na fase de transição entre a terceira e a quarta gerações do jornalismo digital.

Prosseguindo com Barrie Gunter (2003), no seu estudo sobre a evolução dos processos de produção, publicação e entrega de notícias empregando sistemas eletrônicos, as bases de dados são colocadas como agentes fundamentais. Era através delas que tanto as organizações jornalísticas, os profissionais do jornalismo, como também os usuários tinham acesso às notícias por meio dos serviços oferecidos por algumas instituições (tanto do setor de comunicação<sup>102</sup> como nas instâncias governamentais, área médica, do direito e da segurança pública). Inicialmente, o acesso se dava via terminais e, em seguida, já *on-line*, por meio de um computador com *modem*. Para Gunter, desde a década de 70, as bases de dados têm se tornado amplamente adotadas por jornalistas no mundo inteiro.

De acordo com a sua análise, a aquisição e o emprego das BDs no jornalismo seguiu a clássica adoção do modelo da curva do “S”. A afirmação é feita tomando como

---

<sup>102</sup> Conforme ilustrado por Barrie Gunter (ibid, p.21-22), o *Mead Data Central of Columbus*, de Ohio (EUA), desenvolveu, em 1973, o *Lexis*, a primeira base de dados *on-line* disponibilizada totalmente na forma de texto sobre a área do direito. Cinco anos depois, a mesma empresa lançou o *Nexis*, a primeira BD *on-line* contendo publicações nacionais de notícias, como o *The New York Times*, *The Washington Post*, *U.S. News and World Report* e *Associated Press (AP) News Service of Information*. O *Dialog*, maior distribuidor de bases de dados comercial, foi disponibilizado em 1972 e, em 1979, lançou o *National Newspaper Index*, um banco de dados bibliográfico, mais tarde convertido em um serviço de formato textual.

parâmetro estudos realizados nos Estados Unidos. Foi lá que as observações sobre o uso de bases de dados *on-line* indicaram que elas tiveram uma rápida aceitação por parte dos jornais (*early adopters*) e logo se confirmaram como um padrão crescente, alcançando outras empresas de comunicação – as emissoras de televisão e de rádio –, em seguida experimentando fase de consolidação (entre os anos 80 e 90) e depois de estabilização (metade dos 90). Passaram por um período curto de menor pujança, para, depois, experimentar novo crescimento (motivado pela web, a partir de 1995), que impulsionou o aumento gradual na expansão do emprego das bases de dados. Pouco mais de três décadas depois, as bases de dados mantêm seu vigor.

A flexibilidade desta tecnologia é, certamente, um dos seus principais atributos, colocando-a, portanto, como uma solução adequada e fundamental para ser utilizada junto a outros recursos, técnicas e linguagens, que vão estar, preponderantemente, incorporadas no desenvolvimento dos sistemas de gestão de conteúdos de produtos digitais, os conhecidos *Content Management Systems* – *CMS* (Sistemas de Gerenciamento de Conteúdo). Percebendo as bases de dados como uma inovação para o jornalismo e para o jornalismo digital, consideramos que elas reúnem as Características de Inovações elaboradas por Everett Rogers (1995), por possuírem:

- 1) Vantagens para serem adotadas na gestão dos produtos digitais, tanto que o nível de percepção dos benefícios de tal tecnologia tem elevado o número de organizações jornalísticas que a estão empregando em suas operações;

- 2) Compatibilidade com as necessidades requeridas para melhor adequação dos processos e procedimentos quanto à produção de conteúdos mais afinados com os elementos caracterizadores do jornalismo no suporte digital;

3) Complexidade. Ainda que seja considerável o nível de dificuldade, gradualmente se percebe que as organizações jornalísticas, entre outras empresas produtoras de informações noticiosas, vêm migrando para uma plataforma que contempla as bases de dados nos seus produtos digitais;

4) Testabilidade. As BDs são uma tecnologia que pode ser testada e, pouco a pouco – a depender do perfil da organização –, integrada em todas as etapas do processo de produção jornalístico, da organização e qualificação das fontes à apuração, contextualização das matérias e reportagens, edição, até a apresentação;

5) Observabilidade. A visibilidade dos resultados alcançados pelos que cedo aderiram às BDs para estruturar os seus produtos digitais tem feito com que a adoção de tal tecnologia se amplie a cada dia, favorecendo o movimento ascendente e contínuo da curva do “S” para a utilização das BDs no jornalismo;

6) Reinvenção. A flexibilidade como um dos principais atributos das bases de dados tem permitido a esta tecnologia da informação chegar à sua terceira geração de evolução (PIATTINI, 2000; SALGADO, RODRÍGUEZ, 2003) como a mais indicada para a gestão de dados, objetos, gestão do conhecimento, atendendo às necessidades de aplicações para a publicação digital. A reinvenção está também associada aos modos como ela vem sendo utilizada e conceituada na contemporaneidade, notadamente, no reconhecimento das bases de dados como uma forma cultural (MANOVICH, 2001) e, por conseguinte, como um formato (MACHADO, 2004a, 2004b, 2006), um novo paradigma, e com o potencial de estabelecer uma metáfora (BARBOSA, 2004a; 2004b, 2004c, 2005a, 2005b, 2005c, 2006a, 2006b) no jornalismo digital.

## 2.2 EMPREGO DE BASES DE DADOS NAS EMPRESAS INFORMATIVAS

Se a aparição das bases de dados como uma ferramenta para o trabalho jornalístico, na década de 70, representou uma inovação nos modos de obtenção de informação para acrescentar maior contexto e profundidade às notícias e reportagens – além de ter sido uma das primeiras tecnologias empregadas para a entrega eletrônica de conteúdos –, conforme atestam autores como Smith (1980), Caritá (1987), Koch (1991), Pujol (1995), Sellarés (1995), Armañanzas; Noci; Meso (1996), Garrison (1998), Shedden (1998), Paul (1999), Gunter (2003), três décadas depois o emprego das BDs para a gestão de produtos digitais é colocado como o diferencial e a especificidade do jornalismo digital em relação às modalidades tradicionais (COLLE, 2002, 2005a, 2005b; FIDALGO, 2003, 2004, 2005; MACHADO, 2004a, 2004b, 2006; QUADROS, 2004, 2005; LIMA JÚNIOR, 2004, 2005, 2006; BARBOSA, 2004a, 2004b, 2004c, 2005a, 2005b, 2006a, 2006b; HOLOVATY, 2005, 2006a, 2006b).

Não à toa, trabalhos de outros investigadores (HALL, 2001; PAVLIK, 2001, 2005; LAGE, 2001, 2002; PRYOR, 2002; QUINN, 2002, 2005; PALACIOS, 2002a, 2003b, 2003c, 2004; LÓPEZ, GAGO, PEREIRA, 2003; SCHWINGEL; MACHADO; PALACIOS; ROCHA, 2005; SCHWINGEL, 2003, 2004, 2005; GARCÍA et al, 2005; MCADAMS, 2005; SALAVERRÍA, 2005b; CANAVILHAS, 2005, 2006; MIELNICZUK; MARQUES, 2006; GAGO, 2006), ainda que não totalmente voltados para a definição de aspectos relacionados às bases de dados, colocam-nas como um recurso fundamental para o jornalismo digital. Seja como uma tecnologia integrada aos sistemas publicadores de conteúdos, seja na gestão dos produtos, ou mesmo como ferramentas auxiliares, as BDs são descritas como essenciais, notadamente, para que os

*web sites* jornalísticos possam prover conteúdos mais qualificados, tanto no que se refere à composição das peças informativas e sua forma de apresentação, como também quanto à contextualização, isto é, no tratamento mais aprofundado das informações.

Portanto, em consonância com a revisão de literatura, e de acordo com a argumentação desenvolvida nesta tese, defendemos que o emprego de bases de dados no jornalismo sempre foi um condicionante de inovação. Seja atendendo aos propósitos de armazenamento das informações para recuperação e compartilhamento; como fontes de pesquisa e análise, integradas às etapas de apuração e elaboração de textos jornalísticos mais contextualizados; ou usadas para a distribuição de notícias por meio de sistemas como o videotexto, seja para atender às necessidades colocadas para a publicação das edições digitais dos jornais, elas foram importantes protagonistas. Agora, uma vez mais, as BDs emergem na dianteira de um novo processo de inovação.

Ele começou a ser delineado a partir dos primeiros anos do século XXI, através de estudos que passaram a considerar a tecnologia de bases de dados como aspecto diferencial e inovador para o jornalismo no ciberespaço, haja vista as potencialidades nelas existentes para agregar mais qualidade aos produtos jornalísticos. A maneira como muitas empresas informativas têm aderido e se apropriado das BDs, de *softwares* e de técnicas que impulsionam o seu emprego indica que elas seguirão por muito mais décadas como uma tecnologia a garantir inovação para o jornalismo.

Deslindemos, pois, um pouco mais sobre os modos de utilização das bases de dados no jornalismo – tanto naquelas apropriações que as configuram como ferramentas, quanto naquelas relacionadas aos usos e aplicações da tecnologia enquanto tal. Há, de certa maneira, uma vinculação entre elas, ainda que os desdobramentos apontem mesmo para a existência de algumas rupturas quando ao emprego e à conceitualização das BDs.

### 2.2.1 Estrutura para armazenamento de informações

Como mencionado na abertura deste capítulo, o uso de bases de dados no jornalismo esteve, desde o princípio, associado aos procedimentos para armazenamento das informações, com a digitalização dos arquivos. O caso mais emblemático quanto a este modo de emprego surgiu com a efetivação do projeto do *The New York Times Information Bank*, em 1973. Ao lado disso, muitas outras foram as iniciativas de jornais norte-americanos (*The Washington Post*, *Los Angeles Times*<sup>103</sup>, etc.), europeus (*Financial Times*<sup>104</sup>, *The Economist*<sup>105</sup>, entre outros), brasileiros (*O Estado de S.Paulo*, *Folha de S.Paulo*, entre outros) no empreendimento de esforços para digitalizar seus arquivos e, com isso, habilitá-los para uso dentro da redação, bem como disponibilizá-los para acesso ao público externo.

Em geral, os projetos para os bancos de informação jornalística (CARITÁ, 1987) começaram com o arquivamento de resumos, índices e referências para, numa etapa seguinte, oferecer os textos completos. Essa fase ocorre a partir dos anos 80 e se estende ao longo da década de 90 até o momento atual, dependendo do nível de investimento realizado e das prioridades de cada empresa. O *The New York Times*, por exemplo, possui todo o seu arquivo digitalizado, desde 1851 até os dias atuais; assim como o *The Washington Post*, cujo material disponibilizado abrange desde 1877. No Brasil, os arquivos digitalizados de *O Estado de S.Paulo* cobrem o período a partir de 1990,

---

<sup>103</sup> <http://www.latimes.com/>

<sup>104</sup> <http://www.ft.com/home/europe>

<sup>105</sup> <http://www.economist.com/>

porém o jornal está empreendendo esforços no sentido de oferecer todo o seu grande banco de dados histórico, com informações desde o século XIX (SILVA JÚNIOR, 2006, p.289)<sup>106</sup>. Já a *Folha de S.Paulo* disponibiliza o arquivo de 1994 em diante, enquanto o *A Tarde*, desde 1996. O *Correio da Bahia*, por sua vez, mantém *on-line* apenas parte do material de arquivo já digitalizado, cobrindo de 2001 até o presente.

Outra maneira de disponibilização dos arquivos dos jornais se deu através dos distribuidores comerciais de bases de dados *on-line*, a exemplo do *BRS*, que começou vendendo o *Dow-Jones' Information System*, o qual agregava ainda os *abstracts* de notícias do *The Wall Street Journal*. Outros distribuidores, como *Dialog*, *Lexis* e *LexisNexis* (a partir de 1983), também ofereciam bases de dados com arquivos de notícias de muitos jornais, além daquelas implementadas por elas próprias, até hoje figurando como dois dos maiores e principais fornecedores de BDs de informações de diversas áreas.

Estudos sobre o uso dos arquivos nos jornais (SEMONCHE, 1993a, 1993b; PUJOL, 1995; RECIO, 1999; DÍAZ NOCI; MESO, 1999; MACHADO, 2000, 2002) mostram o quanto a possibilidade da digitalização e o uso das redes transformaram as funções dos arquivos dentro do jornalismo. De “arquivos mortos”, antes chamados de “*morgue*”, passaram a ser peças importantes, o que levou as empresas a implementarem os serviços de documentação, mais tarde denominados centros de documentação<sup>107</sup>, com profissionais especializados, para o gerenciamento da sua própria memória, entre outros arquivos de fontes diversas. Além disso, tais departamentos funcionavam como suportes

---

<sup>106</sup> De acordo com Eduardo Mattos, editor-chefe da *Agência Estado*, do *Grupo Estado de S.Paulo*, o grupo trouxe ao Brasil a mesma equipe que fez a digitalização de todo o arquivo do *The New York Times* para realizar um estudo sobre o material do *Estadão*, porém admitiu ser impraticável levar a cabo o projeto neste momento, pois demanda alguns milhões de dólares (ibid.).

<sup>107</sup> A documentação jornalística, segundo Marcos Recio (1999, p.38), se diferencia das demais formas de documentação, como a científica ou a econômica, porque permite um acesso em tempo real aos documentos.

para a elaboração de reportagens investigativas, as quais contavam com o auxílio dos documentalistas, figuras que desempenharam importante papel no incremento da pesquisa *on-line* nas redações, ou, segundo refere Barbara Semonche (1993b), do Jornalismo Assistido por Computador (*CAJ*<sup>108</sup>).

A descoberta do valor dos arquivos para o próprio negócio dos jornais representou um passo fundamental para que, anos mais tarde, eles viessem a integrar as suas edições digitais como um conteúdo essencial e, em paralelo, ser agregado às estratégias comerciais para a geração de recursos, a partir do modelo de negócios adotado. Os arquivos, como afirmado por Marcos Palacios (PALACIOS, 1999, 2002a, 2003b, 2003c, 2004), definem a característica da memória no jornalismo digital como aquela que traz uma das importantes rupturas para esta modalidade em relação às anteriores, pois que ela é o primeiro caso de memória múltipla, instantânea e cumulativa. Conforme explica o autor:

Trabalhando com bancos de dados alojados em máquinas de crescente capacidade de armazenamento e contando com a possibilidade do acesso assíncrono por parte do Utente, bem como de alimentação e Atualização Contínua de tais bancos de dados por parte não só do Produtor, mas também do Utente (Interatividade), o Jornalismo Online, para efeitos práticos, dispõe de espaço virtualmente ilimitado, no que diz respeito à quantidade de informação que pode ser colocada à disposição do seu público-alvo (...) Trata-se da primeira vez que isso ocorre, uma vez que, em todos os suportes anteriores (impresso, rádio, TV), o jornalista convivia com rígidas limitações de espaço (que se traduz em tempo, no caso do rádio e TV). A possibilidade de dispor de espaço ilimitado para a disponibilização do material noticioso é, a nosso ver, a maior Ruptura a ter lugar com o advento da Web como suporte mediático para o jornalismo (PALACIOS, 2002a, p.06-07).

---

<sup>108</sup> A autora emprega *CAJ* em vez de *CAR*.

As agências de notícias, por seu turno, foram e continuam sendo outro destacado setor no emprego das bases de dados, desde a década de 70, para agregar o material produzido anteriormente e a sua organização em formato de arquivos digitais, como assinalado por José Afonso Silva Júnior (2006, p.289). Segundo esclarece, a utilização de bases de dados nas agências de notícias pode ser compreendida em dois níveis: 1) uso interno, atuando na função de tratar a informação e a distribuição para os clientes; e 2) acesso das informações pelos assinantes. Exemplos são o *Reuters Monitor* (sistema que acumulava as cotações e também as notícias publicadas) e o *DataFeature*, da *Associated Press*.

### **2.2.2 Sistemas para entrega eletrônica de notícias**

O termo *on-line* apareceu no meio dos anos 70, sem que lhe dessem muita atenção (GUNTER, 2003, p.18). Na gênese da primeira onda *on-line*, as bases de dados começaram a ser aplicadas para a oferta de grandes volumes de informação, da medicina ao direito, à área da administração pública, até os conteúdos noticiosos de jornais. A entrega eletrônica de notícias e informações se referia, assim, a um conceito amplo de apresentação daqueles conteúdos em uma forma eletrônica a partir da infraestrutura de redes, cuja disponibilização inclui desde CD-ROMs a serviços *on-line* (*Prodigy*, *CompuServe*, *American Online*, entre outros) até a WWW.

Nesta fase, a *Dialog*<sup>109</sup>, pioneira no setor de serviço comercial *on-line* de BDs, destacou-se por ter criado a primeira base de dados comercial, em 1971, batizada como *Pandex* (SUMMIT, 2002; BJORNER; ARDITO, 2003), comercializada a partir de 1972. Um ano depois, a *Lexis* – que se tornaria *LexisNexis*<sup>110</sup>, em 1983 – surgia no cenário dos distribuidores comerciais de bases de dados, oferecendo BDs próprias, além daquelas compiladas a partir do material de jornais, como também fazia a sua concorrente *Dialog*. Desde então, as duas se tornaram líderes globais neste setor, oferecendo bilhões de documentos (textos, imagens, gráficos, entre outros) provenientes de milhares de fontes armazenadas em suas BDs, o que corresponde a mais de 15 terabytes<sup>111</sup> de dados disponíveis.

Em 1980, contudo, foi quando o termo *on-line* tornou-se mais proeminente, principalmente entre as organizações jornalísticas, que passaram a olhar os sistemas eletrônicos como uma ferramenta vital para suas operações e atividades, em parte guiadas pela atualização conferida pelo computador e os *softwares* para facilitar o trabalho diário das suas equipes.

A tecnologia do videotexto é considerada a primeira séria tentativa para a entrega eletrônica de notícias, entre outras informações de serviço (SMITH, 1980; RECODER, ABADAL, CODINA, 1995; SHEDDEN, 1998; SQUIRRA, 1998; GUNTER, 2003), e, com ela, as bases de dados tiveram nova fase de utilização. As experiências com os sistemas de videotexto, ocorridos na segunda metade da década de

---

<sup>109</sup> O criador da empresa, Roger Summit, foi quem desenvolveu o sistema *on-line* para recuperação de informação, em 1967, que impulsionaria, cinco anos depois, o surgimento da *Dialog* como empresa. Em 2000, a *Dialog* tornou-se um negócio da *Thomson Corporation*. A história da *Dialog* está em [www.dialog.com](http://www.dialog.com).

<sup>110</sup> [www.lexisnexis.com](http://www.lexisnexis.com).

<sup>111</sup> Um *terabyte* equivale a 1.000 *gigabytes*. Para se ter uma idéia do que isso pode representar, basta dizer, por exemplo, que um dos HDs (*hardware*) do meu computador, no qual mantenho a maioria dos meus arquivos, possui 35,4 *gigabytes*, e, destes, ocupo atualmente apenas 9,24 *gigabytes*.

70, culminaram com o lançamento comercial do serviço em 1979 e alcançaram expansão principalmente durante os anos 80. Exemplos de sistemas de videotexto são: *Prestel*, *Ceefax* e *Oracle* (Grã-Bretanha); *Keyfax* e os videotextos *Knight-Ridder* e *Times Mirror* (Estados Unidos), *Captain* (Japão), *Telidon* (Canadá), *Ibertex* (Espanha), *Minitel* (França), entre outros. No Brasil, algumas empresas informativas como o *Grupo Estado* e a *RBS*, além das companhias telefônicas, lançaram seus serviços por meio do videotexto. A euforia, assim como os investimentos no videotexto, foi perdendo força a partir do início dos anos 90, quando as atenções se voltaram para a disponibilização, ainda de forma incipiente, de resumos e partes do conteúdo dos jornais impressos, através dos provedores *CompuServe*, *America Online* e *Prodigy*, usando as redes digitais para a circulação.

Em relação ao processo de evolução da publicação eletrônica de notícias e informações, as bases de dados figuram entre os “quatro passos”, conforme citado por Barrie Gunter (2003, p.19), tomando como referência o modelo proposto por F.W. Lancaster, em 1995. Os “quatro passos” são os seguintes: 1) Uso de computadores para gerar publicações convencionais impressas, permitindo a sua oferta de acordo com a demanda ou produzindo publicações personalizadas para as necessidades individuais; 2) Distribuição de texto eletronicamente das versões em papel, que incluía os artigos de texto completo oferecidos pelos distribuidores comerciais de bases de dados; 3) Distribuição na forma eletrônica de publicações impressas, provendo valor agregado, como pesquisa e manipulação dos dados; 4) Criação de publicações que aproveitem as

vantagens das capacidades eletrônicas, como o hipertexto<sup>112</sup>, o som e as imagens em movimento. Ao lado desses “quatro passos”, estão ferramentas, *softwares* e linguagens de programação que permitem a conversão de documentos para um formato que pode ser distribuído para várias plataformas, como o *PDF*, a *Hypertext Markup Language (HMTL)* e a *eXtensible Markup Language (XML)*<sup>113</sup>.

Os desenvolvimentos tecnológicos para a entrega de notícias eletronicamente significaram a criação de um novo mercado de leitores para os jornais e fizeram emergir esta forma de notícia com grande força. Embora as fases de adaptação, desenvolvimento e implementação tenham sido um processo difícil para boa parte das organizações e para os profissionais jornalistas, foram fundamentais para a migração posterior em direção ao ciberespaço, com o estabelecimento das versões digitais na WWW.

### 2.2.3 Fontes para coleta, apuração e recuperação de informações

A informatização das redações, a implementação dos sistemas editoriais, o paralelo processo de digitalização dos arquivos (os bancos de informação jornalística), a oferta de bases de dados comerciais e *on-line* e a criação dos primeiros provedores de

---

<sup>112</sup> Theodor Nelson, autor da expressão, definiu hipertexto em 1974 como uma escrita não-seqüencial, num texto que se bifurca, que permite que o leitor escolha e que se leia melhor numa tela interativa. De acordo com a noção popular, trata-se de uma série de blocos de textos conectados entre si por nexos, que formam diferentes itinerários para o usuário. Sobre o assunto, ver Landow (1997) ou ainda Lévy (1993; 1999). Há autores que distinguem hipertexto como informação puramente escrita, considerando os textos também sonoros, gráficos e visuais como hiperídia (LEÃO, 1999; GOSCIOLA, 2003). Neste trabalho, não fazemos distinção entre hipertexto e hiperídia, podendo o primeiro abarcar não só textos escritos como sonoros, gráficos e visuais (LÉVY, 1993, 1999; MIELNICZUK; PALACIOS, 2001).

<sup>113</sup> Uma das principais características desta linguagem é permitir desvincular a forma do conteúdo. Segundo Sousa (2002, p.07), a *XML* é uma (meta) linguagem de marcação de documentos completamente independente das plataformas *hardware* e *software* que a utilizam, sendo um padrão aberto. A *XML* é a linguagem que torna possível, por exemplo, a troca de dados entre aplicações heterogêneas, fundamentalmente na internet, pois foi projetada para os requerimentos da WWW.

acesso utilizando as redes de computadores conduziram a um panorama de incorporação das BDs ao trabalho diário dos jornalistas. Ainda que o uso ocorresse gradualmente, devido à própria necessidade de aprendizagem e adequação de posturas, o processo de obtenção de notícias, seleção e análise de informações já contava com as bases de dados, efetivamente, entre o final dos anos 70 e os primeiros anos da década de 80.

Dos grandes jornais norte-americanos, esse novo procedimento se espalhou para as redações de jornais de médio e pequeno portes, para as TVs e rádios e, de lá, difundiu-se para outros países (KOCH, 1991; GARRISON, 1998) ainda durante os anos 80, e já com a presença do computador pessoal. O emprego das bases de dados como ferramenta para o trabalho jornalístico ancorava, sobretudo, a elaboração de reportagens investigativas. Podia-se, por exemplo, comparar o número de crimes em diferentes bairros com a estatística demográfica das cidades, ou analisar o acometimento de determinadas doenças em faixas populacionais determinadas, cruzando com taxas de investimento público naquelas áreas e as ações de saneamento e infra-estrutura, etc.

Em 1989, a *Investigative Reporters and Editors (IRE<sup>114</sup>)* abriu o *Missouri Institute for Computer-Assisted Reporting*, na Universidade de Missouri (EUA), onde também funcionava o *National Institute for Computer-Assisted Reporting (NICAR<sup>115</sup>)*. O instituto para a *CAR* surgiu com a missão de ajudar os jornalistas a aprenderem como usar ferramentas de bases de dados para análise de informações visando melhorar suas investigações. A iniciativa foi profícua nas suas ações e realizações e, em 1992, organizou a primeira conferência sobre *Computer Assisted Reporting*. Os objetivos de expansão a partir dali também incluíram a América Latina, com a posterior criação do *Inter-American Institute for Investigative Journalism*, em 1993.

---

<sup>114</sup> <http://www.ire.org>.

<sup>115</sup> <http://www.nicar.org/>.

A contextualização das informações era outro aspecto vinculado à propagação da cultura das bases de dados nas redações. Os repórteres aplicavam técnicas da *CAR*, por exemplo, para provar (ilustrando com gráficos e tabelas) que um terço da população do estado de Nova Jersey vivia em áreas em que a poluição excedia até 20 vezes o limite recomendável (*The Record*<sup>116</sup>); mostraram a evolução da arrecadação de multas de trânsito por diferentes administradores do Alabama, ao longo de dez anos (*The Atlanta Journal-Constitution*<sup>117</sup>); encontraram indícios e provas de fraude eleitoral nas eleições para prefeito de Miami (*The Miami Herald*<sup>118</sup>); provaram que a polícia de Washington era a que mais usava armas e que mais perdia policiais em encontros violentos devido à deficiência de treinamento (*The Washington Post*), entre outras reportagens, (LAGE, 2001, p.162-163), algumas das quais foram premiadas com o *Pulitzer*.

Como afirma Brant Houston (2004), técnicas da *CAR* passaram a ser aplicadas para a reportagem diária, seja para as rondas ou para os grandes projetos que ganham prêmios<sup>119</sup>.

Computer-Assisted Reporting hasn't replaced proven journalistic practices but has become a part of them. It also requires greater responsibility and vigilance. The old standard "Verify, verify, verify", which one learns in basic reporting classes, becomes even

---

<sup>116</sup> <http://www.therecord.com/>.

<sup>117</sup> <http://www.ajc.com/>.

<sup>118</sup> <http://www.miami.com/mld/miamiherald/>.

<sup>119</sup> Atualizando o regulamento daquela que é a mais importante premiação jornalística desde o início do século XX, o *Pulitzer Prizes* estabeleceu regras novas em 2006, permitindo que os participantes concorrentes submetam todo tipo de materiais *on-line*, incluindo bases de dados, gráficos interativos, vídeos, para as diversas categorias. De acordo com Sigg Gissler, administrador do prêmio, "os organizadores acreditam que essa maior incorporação do jornalismo *on-line* reflete o direcionamento dos jornais num mundo midiático em rápida mudança" (Marcos Palacios, em *post* no blog *Jornalismo & Internet*, <http://gjol.blogspot.com>, mantido pelos membros do *Grupo de Pesquisa em Jornalismo Online*, *GJOL*, da FACOM/UFBA, em 29/11/2006).

more critical. “Healthy scepticism” becomes more important, and the practice of interviewing multiple sources and cross-referencing them is more crucial (HOUSTON, 2004, p.03)<sup>120</sup>.

Estava estabelecida, assim, a primeira geração de desenvolvimento da *CAR* ou da Reportagem Assistida por Computador.

#### 2.2.4 A segunda geração da *CAR* e a “*database era*”

A partir de 1995, iniciou-se a segunda geração da *Computer Assisted Reporting*. Foi quando teve lugar um movimento em torno de ferramentas e análises mais sofisticadas, com as combinações e cruzamentos de dados obtidos a partir da web além de outras fontes de referência (GARRISON, 1998, p.08). A web demarcou uma fase fundamental, pois, com o seu surgimento, as empresas fornecedoras de bases de dados migraram seus catálogos de BDs para a WWW, o que facilitou ainda mais a pesquisa, análise, checagem de informações, além, claro, da diversificação de fontes disponíveis. Com a web, outros recursos, como o *e-mail*, o sistema operacional *Windows95*, além de programas como o *PDF*, planilhas eletrônicas de cálculo<sup>121</sup> (*QuattroPro*, *Excell*, *Lotus*

---

<sup>120</sup> “Reportagem Assistida por Computador não tem substituído a prática jornalística comprovada, mas tem se tornado uma parte dela. Ela também requer responsabilidade e vigilância. O velho padrão “Confirmar, confirmar, confirmar”, o qual alguém aprende em classes básicas de reportagem, torna-se mesmo mais crítico. “Ceticismo saudável” torna-se mais importante, e a prática de entrevistar múltiplas fontes e fazer a referência cruzada delas é mais crucial”. (t.A.).

<sup>121</sup> Funcionam como uma calculadora eletrônica mais sofisticada, permitindo efetuar operações matemáticas com diferentes conjuntos de números (por exemplo, transformando valores absolutos em porcentagens) em alguns segundos. Também produzem gráficos que espelham a comparação numérica (LAGE, 2001, p.158).

I-2-3), entre outros *softwares*, vieram somar-se ao conjunto de ferramentas aliadas ao trabalho jornalístico.

Um outro aspecto positivo foi o desenvolvimento de BDs pelos próprios jornalistas, atividade estimulada pelos manuais e cursos oferecidos aos profissionais por entidades como o *NICAR*, e o Instituto para a *CAR*. Matthew M. Reavy (2001, p.134) considera que a incorporação das bases de dados trouxe mais flexibilidade para os jornalistas realizarem seu trabalho, pois, com elas, a elaboração das reportagens podia ser feita de maneira mais rápida e fácil.

Ao organizar informações que vão sendo capturadas regularmente, um repórter pode, por exemplo, montar um banco de dados<sup>122</sup> com informações gerais e específicas sobre um determinado assunto na área de meio ambiente, ou informações de segurança pública, reunindo acontecimentos relativos a crimes contra a mulher, etc. Com as informações organizadas em tabelas<sup>123</sup> (colunas verticais são chamadas “campos”, nos quais pode estar, por exemplo, o nome do autor; e as linhas horizontais, “dados”, nos quais pode estar, por exemplo, o título da obra), o jornalista vai compondo a sua documentação para uso posterior. Aquelas informações poderão ser analisadas com planilhas de cálculo, enquanto outras necessitarão de um software, chamado Sistema Gerenciador de Bases de Dados (SGBD). Estes são empregados para identificar e referenciar qual a informação específica retirar de uma tabela e assim por diante, permitindo fazer cruzamentos e inter-relacionamentos.

---

<sup>122</sup> Entre as BDs mais facilmente disponíveis estão *Access*, que faz parte do pacote *Microsoft Office*, *FoxPro*, *Paradox*, *Approach*, entre outras.

<sup>123</sup> Algumas vezes, uma BD possui uma única tabela de dados (por exemplo, um arquivo contendo nomes e endereços dos contatos das fontes jornalísticas) ou tem duas ou uma coleção de muitas tabelas – neste caso, são bases de dados relacionais – que são planejadas para ser usadas em associação uma com a outra. As BDs relacionais permitem comparações entre duas ou mais tabelas.

Quando as bases de dados alcançaram um alto nível de utilização nas empresas jornalísticas a partir da consolidação da segunda geração da *CAR* – já integrando a *WWW* como a maior das BDs e fonte de informação mais popular nas redações –, Bruce Garrison (1998, p.265) anunciou: “The database era has arrived”<sup>124</sup>.

Oito anos depois, questionado se estaríamos numa nova era de expansão para as bases de dados no jornalismo, ele disse sentir que: “Sim, estamos em uma era das bases de dados no jornalismo, mas elas podem se tornar tão comum que nós não a chamaremos assim no futuro”<sup>125</sup>. Para Garrison, assim como nós tivemos uma vez uma “era do telefone” no jornalismo, após o aparelho ter se tornado uma parte tão comum para o que os jornalistas fazem, ficou desnecessário destacá-lo como tal. Por isso, o mesmo ocorrerá às BDs. “I think computers and databases will evolve to that point in another generation or so of journalists”<sup>126</sup>. E prossegue, assinalando:

Databases require journalists to understand their computers and the software on them. They must learn database tools to use them to their fullest. It requires journalists to develop their quantitative and analytic sides of their minds, also. And it should suggest to most journalists that we look at the bigger picture instead of the close view (GARRISON, 2006)<sup>127</sup>.

---

<sup>124</sup> “A era das bases de dados chegou”. (t.A.).

<sup>125</sup> “I feel we are in a database era in journalism, but it may become so common that we will no longer call it that in the future”. (t.A.). Entrevista concedida à autora, por *e-mail*, enviada em 12/06/2006. [ Mensagem pessoal]. Mensagem recebida por suzana.barbosa.at@gmail.com.

<sup>126</sup> “Acho que os computadores e as bases de dados vão evoluir para aquele ponto em mais ou menos uma geração de jornalistas”. (t.A.).

<sup>127</sup> “Bases de dados requerem que os jornalistas entendam seus computadores e os *softwares*. Eles devem aprender ferramentas de bases de dados para usá-las integralmente. Isso requer também que os jornalistas desenvolvam o lado analítico e quantitativo das suas mentes. E deveria sugerir para a maioria dos jornalistas que nós examinamos a maior imagem em vez de uma visão estreita”. (t.A.).

Bruce Garrison (2006) considera o uso das ferramentas de bases de dados outra forma de obtenção e análise de informação para o jornalismo investigativo. Contudo, apesar dos desenvolvimentos e da cultura que se formou, ele admite que as BDs ainda estão encontrando modos de ser inseridas na rotina jornalística:

We have always used databases in weather and sports journalism, for example, but not as commonly in other subject areas. With the development of small and affordable computing in the past two decades, combined with increasing power of the hardware and software tools, we have begun to move database journalism from a specialization to the mainstream. As more journalists become more computer literate, they will use their computers for more than research online and e-mail. It is only the beginning as I see it. As computing resources improve, so will database journalism. It is driven by the computer industry (GARRISON, 2006)<sup>128</sup>.

Existem inúmeros exemplos de empresas informativas que, atualmente, usam bases de dados efetivamente como um suporte regular para o trabalho jornalístico. Nos Estados Unidos, o próprio Garrison (2006) menciona os mais destacados: *The New York Times*, *U.S. News & World Report*<sup>129</sup>, *The Washington Post*, *Los Angeles Times*, *CharlotteObserver*<sup>130</sup>, *Raleigh News & Observer*<sup>131</sup>, *Miami Herald*<sup>132</sup>, *Cleveland Plain*

<sup>128</sup> “Nós temos usado bases de dados no jornalismo, seja na previsão do tempo ou no esporte, por exemplo, mas não tão comumente em assuntos de outras áreas. Com o desenvolvimento da micro e acessível computação há duas décadas, combinado com o poder crescente do *hardware* e de ferramentas de *software*, nós começamos a mudar o jornalismo de bases de dados da especialização para o *mainstream*. À medida que mais jornalistas se tornem mais habilidosos com o computador, eles usarão seus computadores para fazer mais do que pesquisa *on-line* e correio eletrônico. Isso é apenas o começo de como eu as compreendo. Como os recursos da computação melhoram, também melhorará o jornalismo em bases de dados. Ele é guiado pela indústria da computação”. (t.A.).

<sup>129</sup> <http://www.usnews.com/usnews/home.htm>.

<sup>130</sup> <http://www.charlotte.com/mld/charlotte/>.

<sup>131</sup> <http://www.newsobserver.com/>.

<sup>132</sup> <http://www.miami.com/mld/miamiherald/>.

*Dealer*<sup>133</sup>, *Kansas City Star*<sup>134</sup> e *USA Today*<sup>135</sup>. Outros que podemos citar são o inglês *The Guardian*<sup>136</sup>, os espanhóis *El País*<sup>137</sup> e *El Mundo*<sup>138</sup>, e, no Brasil, *O Globo*, *Folha de S.Paulo*, *Zero Hora*, *O Estado de S.Paulo*.

Para finalizar, no que diz respeito à *Computer Assisted Reporting (CAR)*, o jornalista e professor Philippe Meyer (1999) a considera um conceito obsoleto. Para ele, a *CAR* tornou-se uma lembrança embaraçosa de que o jornalismo estava entrando no século XXI como a única profissão na qual os usuários de computador sentem a necessidade de chamar atenção para si próprios.

After three decades, the time has come to move away from the CAR frame and toward a social science frame. I don't know what label to substitute for CAR. "Precision journalism" sounds dated. We need something that captures the notion in a fresh way (...) The time has come to declare *CAR* victorious and move on to a fresher, more ambitious concept. We need it, and so does the world (MEYER, 1999, p.02)<sup>139</sup>.

Pelo menos por enquanto, ao que parece, ainda não foi encontrado um novo termo e conceito para substituir a *CAR*. Tanto o *NICAR* como o *IRE*, referências

<sup>133</sup> <http://www.plaindealer.com/>.

<sup>134</sup> <http://www.kansascity.com/mld/kansascity/>.

<sup>135</sup> <http://www.usatoday.com/>.

<sup>136</sup> <http://www.guardian.co.uk/>.

<sup>137</sup> <http://www.elpais.com>.

<sup>138</sup> <http://www.elmundo.es>.

<sup>139</sup> “Depois de três décadas, o tempo tem chegado para afastar-se do marco da *CAR* e ir em direção a um marco das ciências sociais. Eu não sei que rótulo substituirá a *CAR*. “Jornalismo de precisão” soa datado. Nós precisamos de alguma coisa que capture a noção em um modo atual. O tempo tem chegado para declarar a *CAR* vitoriosa e mudar para um novo e mais ambicioso conceito. Nós precisamos dele, e deste modo fazer o mundo”. (t.A.).

maiores da Reportagem Assistida por Computador no mundo, continuam a trabalhar com tal denominação. Talvez, o termo, assim como o conceito “jornalismo digital em bases de dados”, possa indicar algum caminho nessa direção.

### 2.2.5 As bases de dados e as primeiras versões de *web sites* jornalísticos

Uma vez tendo diminuído os investimentos dos grupos jornalísticos no sistema videotexto, o movimento nos anos iniciais da década de 90 seguiu na direção dos serviços *on-line* que funcionavam por meio de redes *dial-up* (conexão discada) usando os protocolos para interligação de computadores, segundo a tecnologia internet. Como já mencionado, os mais conhecidos eram *Prodigy*, *CompuServe* e *America Online*. Nesta migração, as bases de dados seguirão novo caminho – o da descentralização dos arquivos com maior propagação da acessibilidade e maiores possibilidades para a sua implementação dentro da oferta informativa dos produtos jornalísticos, inclusive auxiliando na geração de receita para os mesmos.

Em 1992, os exemplos de *sites* de empresas jornalísticas eram os do *St. Louis Dispatch*<sup>140</sup>, *CNN*<sup>141</sup>, *Chicago Tribune* e *Charlotte Observer*. Nos anos seguintes, outros foram surgindo, como *Columbus Dispatch*<sup>142</sup>, *Detroit Free Press*<sup>143</sup> e *The New York Times*, até alcançar um ponto de eclosão em 1995, quando os jornais passaram a se lançar na web por meios próprios, sem a necessidade de recorrer àqueles serviços *on-*

---

<sup>140</sup> <http://www.stltoday.com/>.

<sup>141</sup> <http://edition.cnn.com/>.

<sup>142</sup> <http://www.dispatch.com/>.

<sup>143</sup> <http://www.freep.com/apps/pbcs.dll/frontpage>.

*line*. No Brasil, o processo foi mais lento, e somente em 1995 começou a migração dos jornais para a internet, primeiro via provedores de acesso à rede e, alguns anos depois, já com suas operações independentes.

Foi a partir de 1995 que tiveram início – e prosseguem até hoje – as experimentações com as *breaking news* (últimas notícias), valendo-se da característica da instantaneidade e do dito “tempo real”<sup>144</sup>, com a personalização de conteúdos, com o uso e a integração do áudio, de imagens fixas e em movimento, e mesmo com a busca de um formato para a publicação das informações, usando o *link* como um recurso para estruturar algumas matérias e reportagens, entre outras tentativas de exploração de conteúdos originais destinados especificamente ao suporte digital (ARMAÑANZAS; NOCI; MESO, 1996; PALACIOS; MACHADO, 1997; SHEDDEN, 1998; GARRISON, 1998; SQUIRRA, 1998; PAUL, 1999; HERBERT, 2000; REAVY, 2001; HALL, 2001; WARD, 2002; GUNTER, 2003).

É importante destacar, com Gunter (2003, p.73), que, ao falar de conteúdo eletrônico ou *on-line* dos jornais, também se deve considerar que o formato e a facilidade para o acesso passaram a atrair os leitores não apenas pela qualidade e informação atualizada durante 24 horas. Igualmente relevante foi a nova forma para a apresentação das notícias jornalísticas, incorporando recursos diversos que, assim, as tornariam mais atraentes para o público de leitores e usuários.

Jim Hall (2001) explica que, ao permitir aos leitores acesso direto a bancos de dados de esporte e de estatísticas políticas, uma organização ou operação digital muda radicalmente o relacionamento entre seus jornalistas e os usuários. “O potencial interativo de tal *site* pode ser empregado para permitir aos eleitores encontrarem

---

<sup>144</sup> O *Jornal do Brasil*, primeiro jornal brasileiro na internet, começou a experimentar as *breaking news* em 1996. Os demais *sites*, da *Folha de S.Paulo*, *O Globo*, *Estadão*, também seguiram a tendência. Para uma história do jornalismo digital no país e no exterior, ver: PALACIOS; MACHADO; 1997; MACHADO, 2000; SIMONE, 2001; BARBOSA, 2002).

precisamente as propostas de tributos dos candidatos que podem afetá-los pessoalmente ou para que eles confirmem a posição e performance do seu time e os detalhes históricos” (HALL, 2001, p.52).

Complementando, o professor inglês, da Universidade de Falmouth (Cornuália), vai afirmar ainda que:

Electronic databases mean that we have to add a knowledge of ‘search’ software to the set of skills that we’ve always applied to archives, record-offices and libraries. It also means that there are new ways of telling stories and making them relevant to our readers on an individual basis (...) we, as journalists (or the technical specialists who we work with) would routinely set up search paradigms for our readers to apply to particular databases or sets of databases. That enables them to find out exactly what a story on, say the financing of education, or tax increases, means to them personally. That massively expands the implications of our work (HALL, 2006)<sup>145</sup>.

Por tentativa e erro, acertos, correções de rota, descobertas de modelos de negócios com maior probabilidade para garantir rentabilidade às edições de jornais, além da publicidade, os esforços dirigem-se para o desenvolvimento das operações digitais. Ainda que muitas empresas jornalísticas tenham seguido aquele movimento apenas lançando versões na web para marcar presença sem muito *approach* ou qualquer estratégia definida, a decisão de ingressar no ciberespaço apontava um caminho sem volta, pois os jornais viam declinar a cada dia o número de leitores e precisavam

---

<sup>145</sup> “Bases de dados eletrônicas significa que nós temos que acrescentar um conhecimento de “busca” ao *software* para o conjunto de habilidades que nós sempre temos aplicado aos arquivos, registros oficiais e bibliotecas. E também significa que nós temos novos modos de contar histórias e torná-las relevantes para nossos leitores sobre uma base individual (...) nós, como jornalistas (ou os especialistas técnicos com os quais nós trabalhamos) deveríamos criar rotineiramente paradigmas de busca para nossos leitores para aplicar a bases de dados ou conjuntos de bases de dados. Isso permite a eles descobrir sobre o que trata uma reportagem exatamente, saber sobre o financiamento da educação, ou aumento das taxas, e o que isso significa pessoalmente para eles. Isso expande enormemente as implicações do nosso trabalho”. Entrevista concedida à autora, via *e-mail*, em 18/05/2006. [Mensagem pessoal]. Mensagem recebida por suzana.barbosa.at.gmail.com.

encontrar outras maneiras para atrair o público. Todas as previsões e estudos apontavam a nova mídia internet e a WWW como o lugar para onde tudo convergiria. E assim está sendo.

Com o lançamento dos *sites*, as bases de dados, mais uma vez, representaram – e continuam representando – um aspecto influente, pois, com a descentralização propiciada pela internet, amplia-se o poder dos arquivos para as publicações e, principalmente, torna-se acessível para o público um conjunto de informações de valor precioso para auxiliar nas necessidades de pesquisa de natureza diversa. E, conforme assegura Palacios (2002a, 2003b, 2003c, 2004), a conjugação de memória/arquivo, com as características de instantaneidade, hipertextualidade, interatividade, multimidialidade, e a inexistência de limitações para o armazenamento de informação potencializam de tal forma a memória que é legítimo afirmar que se tem, a partir dessa combinação de características e circunstâncias, uma ruptura com relação aos suportes midiáticos anteriores.

Como forma de jornalismo mais recente, o jornalismo digital – feito no âmbito da internet e que pressupõe a coleta, produção, publicação e disseminação de conteúdos através da web e também de dispositivos móveis como celulares, computadores de mão, *iPods*, etc. – é a modalidade na qual o computador, as bases de dados, entre outras tecnologias, já não são considerados apenas ferramentas, mas, sim, constitutivas de uma prática jornalística.

Em seus dez anos de desenvolvimento, tal modalidade engloba *sites* de diversos tipos: versões de congêneres impressos; *sites* de emissoras de TV e de rádios; produtos eminentemente digitais – nascidos nas redes e para as redes – de portais a portais locais; *sites* agregadores de notícias; *sites* informativos que geram notícias de forma automatizada, estabelecendo o inter-relacionamento de informações e empregando

técnicas diversas para apresentá-las; *sites* de resumos de notícias; *blogs*; *sites* de jornalismo participativo, como *wikinews*<sup>146</sup>, entre outros. Muitos desses produtos foram criados e pertencem não apenas a empresas de comunicação, mas também a empresas de perfis diversos, principalmente do setor de tecnologia, como é o caso do *Google*, com o *Google News*<sup>147</sup>.

No capítulo a seguir, trataremos especificamente sobre **O que é o Jornalismo Digital em Base de Dados (JDBD)**. A partir dos trabalhos de Larry Pryor (2002) e de Luciana Mielniczuk (2003), entendemos que tal modelo emerge numa fase de transição entre a terceira e uma quarta geração de desenvolvimento do jornalismo digital. Neste contexto, acreditamos que a conformação deste paradigma ocorre segundo rupturas, remediações e potencialidades, que despontam com o emprego da tecnologia das bases de dados.

---

<sup>146</sup> [http://en.wikinews.org/wiki/Main\\_Page](http://en.wikinews.org/wiki/Main_Page). Em português: [http://pt.wikinews.org/wiki/P%C3%A1gina\\_principal](http://pt.wikinews.org/wiki/P%C3%A1gina_principal)

<sup>147</sup> <http://news.google.com/?hl=en>



## **O QUE É JORNALISMO DIGITAL EM BASE DE DADOS - (JDBD)?**

## 3

**O que é jornalismo digital em base de dados (JDBD)?**

---

*O jornalismo on-line recorrerá necessariamente à tecnologia das bases de dados como especificidade que o distinguirá substancialmente do jornalismo dos meios tradicionais da imprensa, rádio e televisão.*

- António Fidalgo.

*A plena incorporação pelas organizações jornalísticas da lógica dos Bancos de Dados depende da utilização casada das funções de modelo de estruturação da informação, espaço para criação de narrativas e lugar para ativação da memória.*

- Elias Machado.

*Much of the information that journalists collect, day to day, is structured. Information such as crime reports, obituaries and event listings always follow a certain pattern, which can be richly exploited by databases (...)*

*Information is exponentially more valuable if it's structured.*

- Adrian Holovaty.

### 3.1 UM PARADIGMA NA TRANSIÇÃO DA 3ª PARA A 4ª GERAÇÃO

Como uma atividade criativa, o jornalismo desde há muito tem convivido, periodicamente, com o surgimento de modelos e modos diferenciados para a elaboração de produtos e para a forma como são compostos e apresentados os conteúdos jornalísticos. No processo de expansão desencadeado a partir do século XIX (GIOVANNINI, 1987; BRETON; PROULX, 2000; TRAQUINA, 2002; FRANCISCATO, 2005), a tecnologia sempre foi um fator preponderante para o

aprimoramento dos procedimentos da produção jornalística, do trabalho dos profissionais, da oferta informativa, do formato dos produtos, assim como permitiu vencer distâncias para que a velocidade de circulação das notícias pudesse superar barreiras geográficas e temporais e chegar até o público, satisfazendo as necessidades de informação da sociedade. Ao lado disso, evoluíram também os meios e as diferentes modalidades: da imprensa ao cinema, do rádio à televisão, até à internet (WOLTON, 2000), na qual despontaria a modalidade mais recente – o jornalismo digital.

Em suas etapas de desenvolvimento ao longo de pouco mais de uma década, o jornalismo digital é demarcado por três ciclos ou fases de evolução (PAVLIK, 2001, 2005), que foram denominados como ondas, pelo professor Larry Pryor (2002), e como gerações, pela investigadora Luciana Mielniczuk (2003)<sup>148</sup>. Pela especificidade tratada nesta tese, tomamos como referência o trabalho de Mielniczuk, ainda que incorporando, em paralelo, a caracterização da terceira onda, elaborada por Pryor, para compor a nossa abordagem.

Na classificação formulada por Luciana Mielniczuk (2003, p.31-39) sobre a trajetória dos produtos jornalísticos na web, as três gerações são as seguintes: primeira geração ou fase da transposição; segunda geração ou fase da metáfora e terceira geração ou fase do webjornalismo. Ressaltamos que, embora a autora adote webjornalismo como o termo referente a esta terceira etapa, optamos pela terminologia jornalismo

---

<sup>148</sup> As diferentes terminologias empregadas refletem também o foco trabalhado por cada investigador. John Pavlik (2001, 2005), por exemplo, pensa as fases de evolução do jornalismo na web em relação à produção de conteúdos, percebendo o momento inicial como de reprodução, o segundo com os jornalistas já utilizando recursos como o *link* para compor conteúdos originais e, o terceiro e atual, como aquele momento do jornalismo contextualizado, no qual se experimenta novas formas de *storytelling*. Já Pryor (2002) é mais extenso na sua periodização, abrangendo o processo de disseminação de informações jornalísticas de um modo geral e tendo como marco a tecnologia do videotexto. A primeira onda, neste caso, começaria em 1982; a segunda, em 1993, com os primeiros provedores de acesso disponibilizando informações, e, a terceira, em 2001, marcada pelo maior desenvolvimento, especialização e sofisticação tanto por parte das empresas, da tecnologia disponível, como das equipes profissionais. Sobre o histórico do jornalismo no suporte digital, ver: ARMAÑANZAS; NOCI, MESO (1996); PALACIOS; MACHADO (1997); SHEDDEN (1998); RECIO (1999); MOHERDAUI (2000) e ZAMORA (2002).

digital por considerarmos mais abrangente, englobando os produtos jornalísticos na web, bem como os recursos e tecnologias disponíveis utilizados para a disseminação dessa produção também para dispositivos móveis, como celulares, *iPods*, *MP3 players*, computadores de mão, *smartphones*, etc. O quadro a seguir ilustra melhor cada geração.

Vale notar que essas fases, como assinalado pela autora, não são excludentes entre si, nem são estanques no tempo, ainda que demarquem períodos distintos a partir de 1995 até o momento atual. Em alguns casos, em um mesmo período de tempo, podem-se encontrar publicações jornalísticas para a web que se enquadram em diferentes gerações e, em uma mesma publicação, podem-se encontrar aspectos que remetem a estágios distintos.

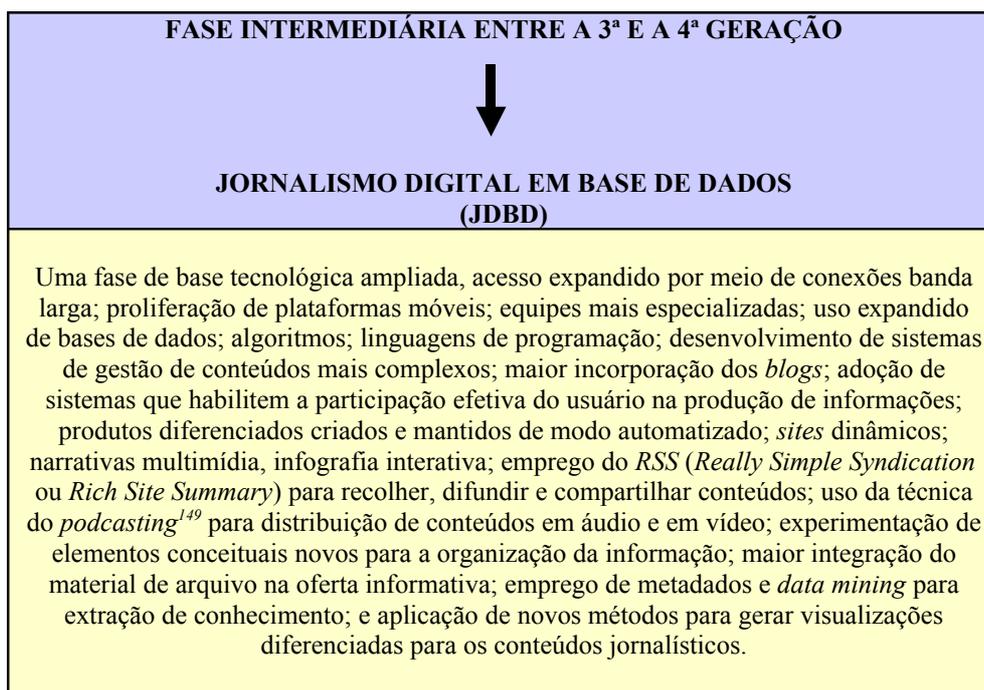
Primeira Geração Transposição	Segunda Geração Metáfora	Terceira Geração Webjornalismo
<p>Produtos oferecidos eram reproduções de partes dos grandes jornais impressos, que passavam a ocupar o espaço na internet. Material era atualizado a cada 24 horas, de acordo com o fechamento das edições do impresso. A rotina de produção também estava atrelada ao modelo estabelecido nos jornais impressos. Não havia qualquer preocupação com uma forma inovadora de apresentação das narrativas jornalísticas. A presença na web correspondia mais à possibilidade de ocupar um espaço, sem explorá-lo, enquanto um meio que apresenta características específicas.</p>	<p>O aperfeiçoamento e o desenvolvimento da estrutura técnica da internet no Brasil (que segue uma tendência mundial, no final dos anos 90 do século XX), permite identificar uma segunda fase. Mesmo atrelados ao modelo do jornal impresso, os produtos começam a explorar as características oferecidas pela rede. É a fase em que o impresso é a referência para a elaboração das interfaces dos produtos. Por isso, é chamada de ‘fase da metáfora’. Recursos que são usados: <i>links</i> com chamadas para notícias de fatos que acontecem no período entre as edições; <i>e-mail</i>, para a comunicação entre jornalista e leitor ou entre os leitores, através de fóruns de debates e a elaboração das notícias passa a explorar os recursos oferecidos pelo hipertexto.</p>	<p>O cenário vai se modificando a partir da crescente popularização do uso da internet e também do surgimento de iniciativas tanto empresariais quanto editoriais destinadas exclusivamente para esse suporte. São <i>sites</i> jornalísticos que extrapolam a idéia de uma versão para a <i>web</i> de um jornal impresso já existente. Nesse estágio, os produtos jornalísticos apresentam recursos em multimídia, como sons e animações, que enriquecem a narrativa jornalística; oferecem recursos de interatividade, como <i>chats</i> com a participação de personalidades públicas, enquetes, fóruns de discussões; disponibilizam opções para a configuração do produto de acordo com interesses pessoais de cada leitor/usuário; apresentam a utilização do hipertexto não apenas como um recurso de organização das informações da edição, mas também começam a empregá-lo na narrativa de fatos.</p>

Quadro 4 – Gerações de desenvolvimento do Jornalismo Digital. Fonte: Mielniczuk, 2003.

A partir deste quadro e agregando as contribuições de Larry Pryor (2002, p.06-07), principalmente naquilo que o autor refere sobre o relacionamento mais próximo entre proprietários e usuários, nas estratégias para atrair a participação dos usuários para a criação de conteúdos, nos investimentos em *softwares* que habilitem poderosas ferramentas de publicação para também permitir a produção de conteúdos originais, uso do vídeo, do áudio, de ilimitadas bases de dados, e na especialização dos profissionais, acreditamos que, neste momento, estaríamos mais apropriadamente numa fase de transição entre a terceira e uma quarta geração do jornalismo digital. Especialmente sobre a participação dos usuários, percebe-se como um ponto característico mais efetivo

na conjuntura atual. Muitos pesquisadores (HALL, 2001, 2005; PRYOR, 2002; MACHADO, 2002, 2003; FIDALGO, 2003, 2004; BOWMAN & WILLIS, 2003; GILLMOR, 2004, 2005; SALAVERRÍA, 2005a, 2005c; ALVES, 2005; PAVLIK, 2005; QUADROS, 2004, 2005; CABRERA, 2005; ROSEN, 2003, 2005a, 2005b, 2006a, 2006b) consideram o jornalismo participativo como um aspecto fundamental que conduzirá o jornalismo digital a uma nova fase de evolução.

Portanto, a etapa intermediária a qual nos referimos parece conduzir a um patamar diferenciado para o desenvolvimento de *web sites* jornalísticos no limiar da segunda década de expansão para essa modalidade, além do que se adequa à perspectiva de criação de uma nova metáfora (BARBOSA, 2004a, 2004b, 2004c, 2005a, 2005b, 2005c, 2006a, 2006b). E um dos aspectos-chave desse estágio, segundo a nossa abordagem, é o emprego da tecnologia de bases de dados para a implementação dos produtos jornalísticos. Com isso, verificamos a emergência do que consideramos como um paradigma novo, isto é, o **Jornalismo Digital em Base de Dados (JDBD)**. Sendo assim, a nossa definição para essa fase de transição, compreende, nomeadamente, a seguinte caracterização, como ilustrado no quadro a seguir:



Quadro 5 – JDBD: novo paradigma surge entre a 3ª e a 4ª geração

A proposição de tal modelo como marco entre a terceira e uma quarta geração do jornalismo digital é justificada, sobretudo, em função da complexificação dos processos para a implementação de produtos jornalísticos no ciberespaço, que vem sendo processada nos últimos três anos. Com a evolução dos recursos, das técnicas e do próprio nível de exigência dos usuários, os produtos necessitaram (e necessitam) adequar melhor seus formatos e a oferta informativa, fazer reformulações para se manterem atualizados e tornar-se, cada vez mais, uma experiência mais autêntica e envolvente para o conjunto dos usuários. Sem contar o fato de o reflexo dessas

<sup>149</sup> A técnica digital chamada *podcasting* é que permite importar da internet um conteúdo de áudio e/ou de vídeo para ouvir/vê-los em *iPods*, por exemplo. Cada usuário pode importar para o seu computador os conteúdos disponíveis, seja uma entrevista, um vídeo de um acontecimento ou outro tipo de conteúdo disponível. Para isso, é necessário instalar um programa agregador, como o *iTunes* (gratuito), com versões para *PC* e *Mac*, que funcionará como receptor de *podcasts*. Outros programas são: *Juice* (*PC* e *Mac*), *Doppler Rádio* (*PC*) e *NewFire* (*Mac*).

reformulações rebater diretamente na adequação das estratégias comerciais e dos modelos de negócio.

Os avanços na tecnologia de bases de dados as tornaram a solução para compatibilizar a incorporação de recursos novos e linguagens de programação para dar forma a produtos dinâmicos e melhor elaborados, a partir do desenvolvimento de sistemas de gestão de conteúdos, visando: aperfeiçoar os sistemas de produção, assegurar maior agilidade, qualidade e descomplicar o trabalho dos jornalistas, entregando sistemas mais fáceis de operar e compatíveis com as características do jornalismo no suporte digital: hipertextualidade<sup>150</sup>, interatividade<sup>151</sup>, multimídia<sup>152</sup>, personalização<sup>153</sup>, memória/arquivo<sup>154</sup>, atualização contínua<sup>155</sup> (PALACIOS, 1999,

---

<sup>150</sup> Uso do recurso do *link* para organizar a publicação e para a construção das narrativas jornalísticas em camadas ou blocos. Uma das características intrínsecas à própria web e fundamental para o jornalismo digital. Os interrelacionamentos, cruzamentos, a flexibilidade combinatória entre conteúdos são possibilitados pelo emprego do *link* na composição das peças informativas.

<sup>151</sup> Incorporação do *e-mail*, *chats*, fóruns para estabelecer relação entre a publicação e o usuário, entre o usuário e os jornalistas e entre os próprios usuários. A interatividade é outra condição essencial para a implementação de conteúdos no jornalismo, e favorece que os usuários atuem como produtores de informação, segundo níveis distintos de participação e de acordo com a linha editorial e princípios definidos por cada publicação.

<sup>152</sup> Consiste no emprego do áudio, imagens fixas e em movimento, infográficos, entre outros, para apresentação de conteúdos. Tais recursos podem estar integrados numa mesma peça informativa ou serem disponibilizados separadamente, por justaposição (SALAVERRÍA, 2005a).

<sup>153</sup> Consiste em adequar os conteúdos de uma publicação aos desejos ou necessidades pessoais de grupos específicos de usuários ou a cada um deles, de modo particular.

<sup>154</sup> Um dos elementos primordiais, pois é com a internet que o jornalismo se liberta dos limites de tempo e espaço para explorar os conteúdos, principalmente, aquele material armazenado, ou seja, o arquivo. Como já mencionado no capítulo anterior (e posteriormente ainda se vai explorar), a característica da memória representa uma ruptura (PALACIOS, 1999; 2002a; 2003b; 2004) em relação aos modos como se trabalha com os arquivos nas modalidades tradicionais. No jornalismo digital, tem-se o primeiro caso de memória múltipla, instantânea e cumulativa.

<sup>155</sup> Característica que assinala a quebra da periodicidade. A atualização das informações jornalísticas pode ocorrer durante 24 horas, 365 dias do ano. A descentralidade da rede, e os sistemas de gestão de conteúdos permitem que os jornalistas possam trabalhar, enviar e publicar informações de qualquer lugar, assim como podem atualizar *blogs*, sejam eles pessoais ou mantidos nos *sites* das próprias empresas de comunicação.

2002a, 2003b, 2004; BARDOEL & DEUZE, 2000; ZAMORRA, 2002; MIELNICZUK, 2003; SALAVERRÍA, 2005b).

Ao propormos o modelo **JDBD** como paradigma de uma fase de transição, ressaltamos que o quarto estágio se encontra ainda em gestação. Porém, seja qual for a configuração que irá consolidá-lo, as bases de dados serão um elemento estruturante, pela flexibilidade intrínseca a elas, assim como o potencial de reinvenção que elas possuem. Carla Schwingel (2005), num artigo em que especula sobre as tendências que conduziriam à essa quarta geração do jornalismo digital, coloca as BDs como um componente fundamental, que, associadas a sistemas de gestão de conteúdos automatizados para a apuração, edição e veiculação de informações, impulsionariam a “efetiva industrialização de processos que hoje são elaborados muitas vezes de forma intuitiva e artesanal” (2005, p.01).

Assim, apresentamos neste capítulo o **Jornalismo Digital em Base de Dados (JDBD)**. Concentramo-nos na fundamentação teórica, cuja elaboração tem como referência contribuições de estudos advindos da área da nova mídia, do jornalismo e do jornalismo digital, e das ciências da computação. Em paralelo, buscamos ilustrar os conceitos através do material empírico que integra o conjunto de casos da amostra trabalhada. Tal aporte teórico nos leva a definir o **JDBD** como um paradigma, e, assim, no capítulo quatro, apresentamos o conceito que colocamos em operação, identificando as funcionalidades das bases de dados para o jornalismo digital, assim como as categorias que lhe conferem especificidade.

### 3.1.1 Atribuição de um novo *status*: a remediação e a extensão do conceito

Certamente, alguém há de questionar: como uma tecnologia que desde os anos 70 auxilia o trabalho jornalístico pode agora estabelecer um paradigma no jornalismo digital? Para responder a tal indagação, fomos buscar um conceito do campo das novas mídias. Trata-se de *remediation* ou remediação (BOLTER & GRUSIN, 2000), considerada pelos autores norte-americanos como a palavra-chave para a mídia digital, na qual os meios anteriores são remediados. *Remediation* é a representação de um meio em outro, implicando no reconhecimento do meio anterior, da sua linguagem e da sua representação social para se estabelecer em um novo suporte. Com isso, Jay David Bolter e Richard Grusin (2000) querem significar que todos os meios passam por processos de remediação, e melhorias, uma vez que têm o seu sistema de produção afetado pela chamada nova mídia.

(...) As a digital network, cyberspace remediates the electric communications networks of the past 150 years, the telegraph and the telephone; as a virtual reality, it remediates the visual spaces of painting, film, and television; as a social space, it remediates such historical places as cities and parks and such nonplaces as theme parks and shopping malls. Like other contemporary mediated spaces, cyberspace refashions and extends earlier media, which are themselves embedded in material and social environments (BOLTER & GRUSIN, 1999, p.182)<sup>156</sup>.

---

<sup>156</sup> “Como uma rede digital, o ciberespaço remedeia as redes de comunicação elétrica de 150 anos atrás, o telégrafo e o telefone; como uma realidade virtual, ele remedeia os espaços visuais da pintura, do filme, e da televisão; como um espaço social, ele remedeia deste modo lugares históricos como cidades e parques e não lugares tais como parques temáticos e *shopping centers*. Como outros espaços contemporâneos mediados, o ciberespaço reconfigura e estende os meios prévios, os quais estão enraizados em ambientes materiais e sociais”. (t.A.).

De acordo com os autores, as novas mídias remedeiam e, nesse processo, melhoram seus predecessores (o jornal, a revista, o rádio, o telefone, a arte, os modos de publicar, a experiência social, o espaço urbano). A internet, por sua vez, remedeia todos os meios, aperfeiçoando-os em muitos aspectos e acrescentando recursos novos, enquanto a web, especificamente, tem uma natureza remediadora, operando de modo híbrido e inclusivo. “A web remedeia os jornais, as revistas e a publicidade gráfica, tornando as revistas da internet (ou revistas digitais) e os serviços de notícia gêneros importantes e populares”, afirmam Bolter e Grusin (2000, p.198).

Tal conceito se mostra apropriado para a fundamentação da nossa hipótese, na medida em que nos permite perceber a ampliação do significado de bases de dados, compreendendo a sua concepção tanto como forma cultural simbólica (MANOVICH, 2001), quanto como a de formato (Machado, 2004a, 2004b, 2006), metáfora (BARBOSA, 2004a, 2004b, 2004c, 2005a, 2005b, 2005c, 2006a, 2006b) e paradigma no jornalismo digital. Ademais, está em sintonia também com o que denominamos de renovação do movimento ascendente da curva do “S” para a adoção de bases de dados, devido ao seu potencial para a gestão de produtos jornalísticos digitais.

As bases de dados, mesmo não sendo um meio de comunicação, um espaço visual, social ou urbano, são remediadas pelo fato de que a internet vai garantir novas técnicas e linguagens para a sua construção e aplicação (SOUSA, 2002). Elas sofrerão remediações e possibilitarão algumas rupturas, ganhando novas funcionalidades de acordo com os usos e apropriações no jornalismo digital. Além disso, como observado pelos próprios Bolter e Grusin (2000), foram os bancos de dados textuais que precederam a introdução da web e dependeram dos serviços anteriores da internet ou mesmo da gravação digital para realizar a remediação do livro (BOLTER & GRUSIN, 2000, p.201).

Por um lado, o emprego e a utilização de BDs no jornalismo digital opera remediações nos sistemas de produção, de obtenção e de estruturação das informações jornalísticas, no âmbito da composição, da apresentação e da visualização dos conteúdos, e, em outro, vai gerar inovações quanto aos modos de fazer jornalismo nas redes digitais, configurando, então, um cenário de dupla via caracterizado por remediações e rupturas (BARBOSA, 2004b, 2005a, 2005b, 2006a, 2006b).

Expomos tal argumento a um dos autores – Jay David Bolter<sup>157</sup> –, que se disse surpreso com a conexão entre o conceito de *remediation* e a noção de bases de dados tal como empregamos na nossa investigação. Porém, considerou: “The notion of greater authenticity does fit in with remediation”<sup>158</sup>. No seu comentário, Bolter explica:

To me, an analysis based on remediation would look at the relationship between online journalism as an attempt to imitate and surpass earlier forms (...) The database argument is based, I think, on the notion that a rupture (as you suggest) has occurred. Now the journalist and reader have a database of materials available and can fashion their own paper. On the other hand, perhaps the database could be understood as the feature that makes the online journal a more authentic or compelling experience in comparison the printed newspaper. The notion of greater authenticity does fit in with remediation of course (BOLTER, 2005)<sup>159</sup>.

---

<sup>157</sup> A oportunidade para expor a nossa apropriação do conceito ocorreu na palestra realizada pelo autor no dia 06/06/2005, na Universidade Lusófona de Humanidades e Tecnologias (ULHT), em Lisboa, Portugal, e seus comentários aqui reproduzidos foram enviados por *e-mail* no dia 21 de Maio de 2005.

<sup>158</sup> “A noção de maior autenticidade encaixa com a de remediação”. (t.A.).

<sup>159</sup> “Para mim, uma análise baseada em remediação olharia para a relação entre jornalismo *on-line* como uma tentativa de imitar e superar formas anteriores (...) O argumento sobre as base de dados é apoiado, penso eu, na noção de que uma ruptura (como você sugere) ocorreu. Agora o jornalista e o leitor têm uma base de dados dos materiais disponíveis e podem moldar seu próprio jornal. Por outro lado, talvez as bases de dados pudessem ser compreendidas como a característica que torna o jornal *on-line* uma experiência mais autêntica ou envolvente para o usuário, em comparação com o jornal impresso. A noção de maior autenticidade encaixa com a de remediação, é claro”. (t.A.).

As rupturas estariam na quebra de um certo padrão até então empregado para a organização e a estruturação das informações, para a construção das peças informativas, para o arquivamento e recuperação dos conteúdos publicados, como também para o uso de elementos conceituais novos e métodos diferenciados para a apresentação das informações. Mas não só. Ao nos apropriarmos do conceito de *remediation* (BOLTER & GRUSIN, 2000), o objetivo é mostrar como as bases de dados podem ter seu emprego expandido em função dos desenvolvimentos tecnológicos, em relação ao que até então se fazia e se conhecia, ainda que partilhando de usos e procedimentos anteriores em relação à sua adoção.

Conforme demonstrado pelos autores norte-americanos, todas as inovações trazidas pelas novas mídias, cuja característica definidora é a matriz digital, não ocorrem de maneira divorciada dos meios existentes. Ao contrário, as novas mídias remedeiam, aperfeiçoam seus predecessores, tanto o jornal, a revista, o rádio, o telefone, a arte, o vídeo, a fotografia, a comunicação face a face, os modos de publicar, assim como a experiência social e o espaço urbano<sup>160</sup>.

Há, no entanto, quem refute o conceito de *remediation*<sup>161</sup>. Entre eles está Geert Lovink (2003, p.152-153), para quem os novos meios abrem possibilidades para outras formas de narração, assim como para novas estéticas. Para o defensor da cultura crítica

---

<sup>160</sup> Sob esse aspecto, o conceito de remediação é apropriado por Stephen Graham (2004) para analisar as remediações que as novas mídias operam na vida urbana, a partir de seis pontos. Entre eles, sob o ponto de vista das pressões de continuidades e discontinuidades, Graham diz que a tão chamada “era da informação” é melhor considerada não como uma revolução, mas como um complexo e sutil amálgama de novas tecnologias e meios fundidos, e remediando os antigos (GRAHAM, 2004, p.18-19). Já Timothy W. Luke (2002) aplica o conceito de remediação para repensar a cultura política e o poder na sociedade contemporânea em função das TICs e, em especial, da internet.

<sup>161</sup> Lev Manovich (2006, p.04-07), ao comentar sobre as contribuições de Alan Kay e seu grupo de pesquisa na Xerox Parc, nos anos 70, afirma que foi lá o lugar onde a identidade do computador como uma máquina de remediação (*remediation machine*) teve lugar, sendo transformada depois em uma *universal media machine*. Numa crítica ao conceito de *remediation*, Manovich vai dizer que se prestarmos atenção para além do que está na aparência é possível perceber novas funções verificadas num meio, mas que não estavam presentes previamente. Um exemplo seria a fotografia digital, que permite ao usuário editá-la, modificá-la, combiná-la e movê-la para outros dispositivos, em modos distintos daquilo que se podia fazer com o seu equivalente tradicional.

da internet, se a rede tem sido tomada como algo verdadeiramente novo – como todos os visionários reivindicam –, o *streaming* poderia ser o primeiro exemplo de como suplantar o antigo modelo um-todos do meio *broadcasting*. O *streaming media*, afirma Lovink, explora novas condições de rede e não busca reconstruir o velho mundo do audiovisual no suporte digital.

Por isso mesmo ele vai dizer que o *streaming* resiste à remediação, pois só se pode falar de *streaming media* onde existem canais abertos para *feedback*. Porém, se fôssemos analisar o *streaming media* – o que não é nosso objetivo –, poderíamos descobrir que, ainda que ele se torne possível com a internet, o que o precede certamente é o meio audiovisual. Ou seja, seria possível também perceber algum tipo de remediação para esta tecnologia.

Portanto, na nossa perspectiva e tendo em vista nosso objeto, remediação revela-se uma idéia adequada para ajudar a construir formulações teóricas em torno das especificidades das bases de dados no jornalismo digital, pois contribui para verificar as rupturas processadas em relação a padrões anteriores e, ainda, permite identificar as potencialidades de tal tecnologia para o jornalismo. Sendo assim, passaremos ao exame de algumas proposições que corroboram a ampliação do significado das bases de dados e identificam os usos e as potencialidades para o seu emprego no jornalismo digital.

### **3.1.2 Aporte teórico e o estado da arte das bases de dados no jornalismo**

A partir dos primeiros anos do século XXI, localizamos os trabalhos de investigadores que têm ampliado a dimensão conceitual para as bases de dados no jornalismo, em especial no jornalismo digital. O aporte teórico desenvolvido é

sistematizado tendo-se como referência empírica os produtos digitais, obviamente, mas, sobretudo, ele é inspirado pela percepção de um cenário mais amplo, agregando aí também as experimentações recentes que vêm sendo realizadas em diversos países.

Raymond Colle, em seu livro *Explotar la información noticiosa. Data mining aplicado a la documentación periodística*, de 2002, talvez seja um dos pioneiros em apresentar e discutir as vantagens das bases de dados para o chamado “diário electrónico” para o acesso personalizado dos usuários aos sistemas documentais dos jornais. Além disso, verificou que também o jornalismo poderia se beneficiar da técnica do *data mining* ou mineração de dados<sup>162</sup> e do KDD (*Knowledge Discovery in Databases* ou prospecção de conhecimento embutido nos dados<sup>163</sup>) para “descobrir os fios que unem os fatos e explicar os êxitos e fracassos em múltiplos campos” (COLLE, 2002, p.19).

Para o autor belga radicado no Chile, o jornalismo, como grande consumidor e difusor de informações, a partir dali, não poderia ignorar a importância de técnicas como o *data mining* para o seu trabalho. E, então, empreendeu um detalhado estudo sobre o *data mining* aplicado ao jornalismo, destrinchando termos e explicando os processos para se realizar a mineração de dados, descobrindo informações úteis escondidas por entre os registros inseridos numa base de dados jornalística e demonstrando os modos como aquele conteúdo poderia ser visualizado. A parte empírica do trabalho foi feita com a revista online *Temas Digitales de la Comunicación*

---

<sup>162</sup> Ou, ainda, exploração de dados. Conjunto de operações de diversos tipos que se pode realizar para explorar relações, tendências e projeções em conjuntos de dados de grandes dimensões (COLLE, 2002, p.255). Como nos referimos no primeiro capítulo, *data mining* é uma forma automática de aquisição de conhecimento, uma técnica que torna um sistema inteligente à medida que auxilia no processo de decisão, sem a interferência humana. Apesar disso, como explica Navega (2002), esse processo que se vale de algoritmos para processar os dados e encontrar padrões novos e valiosos, ainda requer forte interação com os analistas humanos, pois, são eles os principais responsáveis pela determinação do valor dos padrões encontrados. As ferramentas de *data mining* se baseiam no conhecimento acumulado que está invisível nas bases de dados.

<sup>163</sup> É um processo que engloba a mineração.

(TDC), criada por ele no âmbito do Centro de Estudios Mediales da Universidad Diego Portales.



Figura 20 - Tela da *home* de artigos da TDC.

Fonte: Imagem cedida por Raymond Colle. Out.2005

A TDC, que esteve disponível entre agosto de 2000 e janeiro de 2005, era um produto de caráter dinâmico, construído sobre bases de dados relacionais e com sistema de metainformação<sup>164</sup> (estatísticas *on-line*). Segundo Raymond Colle, a TDC acumulou mais de nove mil notícias sobre tecnologias digitais de comunicação, que gerou uma base de dados com material divulgado por aproximadamente 70 fontes (entre elas, *Noticias.com*, *Wired*, *CNN* e *El Mercurio*): “Pero no he consolidado aún ni re-publicado la meta-información correspondiente a ese total (que estuvo en línea en enero pasado).

<sup>164</sup> Informação acerca de um conjunto de informações, por exemplo, a temática que têm em comum (COLLE, 2005a). Segundo o autor, este conceito se utiliza essencialmente para designar conhecimentos associados a conjuntos de dados e, no campo da informática, remetem em particular ao conhecimento relacionado com a estrutura e os conteúdos das bases de dados. Existem dois tipos de metainformação: - a que é prévia ao processamento e consiste em todo o conhecimento que se tenha da estrutura dos mesmos. É em função deste conhecimento que se realiza o desenho lógico (conceitual) da base de dados e das tabelas (desenho físico) que a conformam; - e a que é posterior ao processamento, ou seja, quando já se tem os dados acumulados e, ao lê-los em uma determinada ordem, se pode perceber o número de vezes em que o nome de um lugar aparece, por exemplo. O produto digital metainformativo, diz Colle, é aquele que se estrutura a partir de uma base de dados e oferece aos usuários não apenas um sistema de busca refinado, mas também serviços complementares como estatísticas dinâmicas *on-line* (que se ajustam a cada instante, à medida em que se ampliam os conteúdos da base de dados).

Aunque cuento con los archivos que me permitirían hacerlo”, afirmou<sup>165</sup>. O último estudo sobre a *TDC* analisou as cinco mil notícias, marca atingida em novembro de 2002. Atualmente, o arquivo da revista está disponível no *web site* do professor, no qual ele também mantém a *Notibits*<sup>166</sup>, lançada em 2005, como uma espécie de continuação privada da revista anterior. No entanto, ela ainda não está operando sobre bases de dados dinâmicas nem com sistema de metainformação. Ou seja, embora rode sobre base de dados, Colle não a considera uma publicação dinâmica em essência, porque não agrega ainda aqueles recursos implementados na *TDC*.

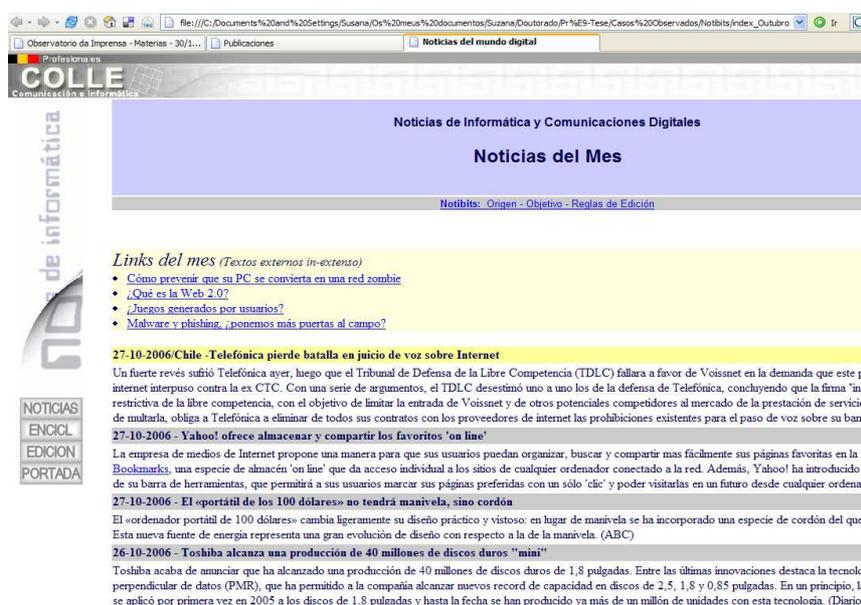


Figura 21 - Revista *Notibits de Informática Y Comunicaciones Digitales* ainda funciona sem sistema de metainformação. Out.2006.

António Fidalgo (2003), por sua vez, foi um dos primeiros a analisar a especificidade conferida pelas bases de dados para a estruturação e a organização das

<sup>165</sup> “Mas ele não está consolidado nem ainda se republicou a metainformação correspondente a esse total (que esteve *on-line* até janeiro passado). Ainda conto com os arquivos que me permitiriam fazê-lo”. Informação concedida à autora, via *e-mail*, em 15/08/2005. Ao que parece, pelo acompanhamento que temos realizado, o trabalho ainda não foi feito.

<sup>166</sup> [www.colle.cl/notibits](http://www.colle.cl/notibits).

informações num produto digital, percebendo a flexibilidade para a forma de apresentação das notícias (e dos produtos, em última instância), assim como as vantagens fornecidas pelas BDs para assegurar objetividade às mesmas. No artigo sobre a sintaxe e a semântica das notícias num produto estruturado em bases de dados, o catedrático português discorre sobre: o carácter distintivo de um produto criado em HTML e outro em bases de dados; além de introduzir a noção de arquivo *on-line* da coleção<sup>167</sup>, aquele material constituído pelas edições publicadas por um jornal e que pode ser agregado ao produto digital para estabelecer os elos de uma cadeia de acontecimentos. Deste modo, diz ele, o passado pode condicionar o presente, na justa proporção em que pode ser recuperado e, de novo, estar presente à atenção.

Para Fidalgo, a grande diferença entre um produto em HTML e outro em bases de dados é que, no primeiro caso (HTML), ele é um produto único, ainda que recorrendo a *templates*, ou seja, constituído por páginas estáticas, sempre com o mesmo conteúdo<sup>168</sup>. Já num produto estruturado sobre uma base de dados, o conteúdo é dinâmico e será sempre o resultado de uma pesquisa (*query*) nas notícias inseridas e da estrutura da base de dados<sup>169</sup>. De modo mais detalhado, António Fidalgo assinala:

---

<sup>167</sup> Demarca, deste modo, a distinção em relação ao arquivo do jornal, que é muito mais vasto e inclui todo o material: texto, fotografias, etc, publicado ou não, que fornece o fundamento histórico do jornal.

<sup>168</sup> Por exemplo, em um site estático em html, para representar visualmente 5.000 notícias, seriam necessárias 5.000 páginas diferentes. Segundo Fidalgo (ibid.:, p.51), para alguém que saiba um pouco de hipertexto e de bases de dados *on-line*, será fácil verificar, consultando a barra de endereços, ou seja, através da forma como se transita de *url* para *url* dentro do mesmo jornal, se se trata de um produto feito só em *html* ou se corre sobre uma BD.

<sup>169</sup> Os *web sites* dinâmicos contam com uma, duas, ou três páginas físicas para representar visualmente, por exemplo, 10 ou 15 mil notícias, pois são flexíveis e a informação pode ser representada a partir de um sistema que negocia com o servidor e o usuário (LÓPEZ, GAGO, PEREIRA, 2003, p.197). Outra diferença, como explica Mielniczuk e Marques (2006, p.02), é que os *sites* dinâmicos utilizam linguagem de programação, o que permite a modificação automática, sem a intervenção manual de um programador. As bases de dados são o fator primordial, nesse caso, pois são elas que possibilitam a construção de *sites* dinâmicos, aliando linguagem de programação, entre outros recursos.

(...) Os jornais assentes em bases de dados distinguem-se entre os jornais on-line por não terem edições fixas, pelo facto de uma edição ser apenas uma configuração possível gerada pela base de dados. Tal configuração, porém, é automática (...) É a estrutura da base de dados que determina a forma como as diferentes notícias aparecem conjugadas na apresentação on-line. Existe, por conseguinte, um rigor sintáctico na organização das notícias, que é ao mesmo tempo consequência e tradução da estrutura lógica da base de dados (...) Compreender-se-á a função sintáctica da base de dados, através das diferentes funções que assume na feitura do jornal, desde a produção até à edição e apresentação das notícias (FIDALGO, 2003, p.55-56).

Outro importante conceito introduzido por Fidalgo é o de resolução semântica, que foi melhor desenvolvido num trabalho posterior (2004). Como explica o autor, tal como uma imagem digital aumenta a sua qualidade com o aumento da resolução gráfica (o número de *pixels* por centímetro quadrado), também a pluralidade e a diversidade das notícias *on-line* sobre um evento aumentam a informação sobre o mesmo, aumentando, assim, a resolução semântica (FIDALGO, 2004, p.02):

Uma primeira notícia sobre um acontecimento, que à partida surge com um determinado sentido, pode ser complementada, alterada, corrigida, à medida que outras notícias sobre o mesmo acontecimento se lhe seguem. O que, de início, tinha contornos indefinidos, deixando múltiplas hipóteses em aberto, vai ganhando sucessivamente formas cada vez mais definidas (...) À medida que forem chegando notícias subseqüentes, a notícia do que ocorreu vai ganhando forma, ou seja, aumenta a sua resolução semântica (FIDALGO, 2004, p.03).

Por exemplo, uma informação ao ser publicada no canal de “últimas notícias” de um produto digital aparece, inicialmente, com uma baixa resolução. Contudo, a partir da apuração e da contextualização do acontecimento, a densidade semântica vai aumentando progressivamente. Se considerarmos a participação dos usuários, acrescentando comentários, complementos à informação, críticas e sugestões, bem como a inserção de áudios de entrevistas, imagens fixas e em movimento, e

infográficos, teremos um aumento contínuo da resolução semântica, cuja meta a atingir seria o estado em que todas as informações sobre o evento estariam disponíveis.

Uma vez disponíveis em sua plenitude, as informações serão lidas e consultadas de modo simultâneo ou não, dado o contexto policrônico ou multitemporal (SALAVERRÍA, 2005b, p.23) que caracteriza o ciberespaço. Deste modo, a resolução semântica estaria assegurada também no arquivamento e recuperação dos conteúdos. Prosseguindo com António Fidalgo (2004), este vai dizer que o conceito de resolução semântica só faz sentido se os elementos informativos sucessivos e progressivos seguirem uma ordenação. Isso ocorre no processo de classificação interna que, por sua vez, proporcionará mudanças também na classificação externa, ou seja, na forma como as notícias (e os conteúdos jornalísticos de modo geral) serão apresentadas. Outra ressalva feita refere-se à estruturação da própria notícia, e, sobre este aspecto, ele diz: o contínuo da informação *on-line* não se adequa ao formato de pirâmide invertida.

Como afirma o pesquisador, a feitura de uma notícia *on-line*, mediante uma base de dados, apesar de responder às célebres perguntas: O Que, Quem, Quando, Onde, Porque e Como, pode fazê-lo de um modo diferente (FIDALGO, 2004, p. 05). Qual seja: as notícias num produto digital são dadas de forma lacunar, deficiente, num primeiro momento, para, em seguida, ser complementadas, modificadas e até corrigidas. Por isso, a imediatividade, muitas vezes, sobrepõe-se às exigências da objetividade e da verificabilidade.

Assim, a definição dos campos de classificação interna para as notícias (quanto à autoria, tipo de evento, grupo social, faixa etária, situação econômica, etc.) ordenará o acréscimo e a distribuição da informação, assim como a progressiva precisão para assegurar a resolução semântica. Esta será maior ou menor, mas a tendência é aumentar no jornalismo assente em bases de dados, pois nele, conforme atesta António Fidalgo, a

resolução semântica aparece consubstanciada na própria notícia, na sua apresentação *on-line*, pois que é apenas uma descrição dos acontecimentos que vai sendo sucessivamente pormenorizada, complementada e corrigida (FIDALGO, 2004, p.07).

António Fidalgo, como diretor do Laboratório de Comunicação On-line (Labcom) da Universidade da Beira Interior (UBI, Portugal), foi idealizador do *Akademia*, um produto de conteúdo académico, que esteve ativo de outubro de 2004 até os primeiros meses de 2005. O objetivo do *Akademia* era integrar todas as notícias numa base de dados, organizada por múltiplas entradas, datas, seções, lugares, intervenientes, tipos de fatos noticiosos, etc., incorporando o áudio, o vídeo, o *link* para construir as peças informativas.

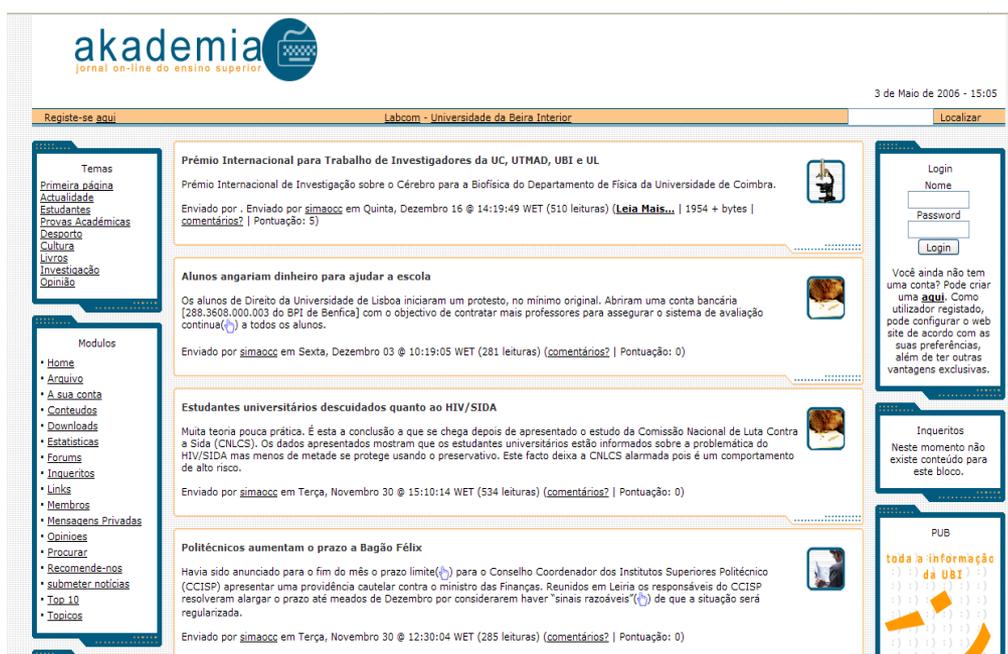


Figura 22 - O *Akademia* foi o primeiro produto do *Labcom* estruturado em base de dados. Maio.2006.

João Simão (2005), jornalista responsável pela atualização das informações do *site*, reporta que o *Akademia* apresentava uma estrutura-base com as notícias publicadas em blocos no centro da tela, identificadas com uma pequena imagem (ícone) que as

inseria dentro de um dos temas da oferta informativa. Apesar das idealizações iniciais para o produto, ele se revela um tanto quanto inconformado, pois declara: “O *Akademia* mais não fez que usar as convenções típicas de estrutura, navegação e funcionamento de um *site*”. Contudo, Simão diz que a grande inovação do produto podia ser vista nos blocos de notícias e na forma como a notícia era escrita<sup>170</sup>.

De certa forma, parte do conteúdo do *Akademia* está integrado em uma das seções temáticas (“Ensino superior”) do *Urbi et Orbi*<sup>171</sup> – jornal laboratório do curso de Ciências da Comunicação da UBI, criado em 2000, também idealizado e dirigido pelo professor António Fidalgo – que foi reformulado entre janeiro e maio de 2006 e passou a ser estruturado em bases de dados. O novo *Urbi* foi lançado em Junho de 2006, e, de acordo com João Canavilhas, um dos diretores-adjuntos do jornal e coordenador do projeto de reformulação, a decisão de passar a publicar o *site* estruturando-o em bases de dados, deveu-se, primeiro, ao fato de que a “UBI tinha e tem responsabilidade no lançamento de novos modelos, de inovar”<sup>172</sup>.

Ele explica: “Identificamos que o desenvolvimento do jornal havia estagnado e já não se verificava evolução no conteúdo. Então, a idéia de mudar surgiu a partir da investigação e porque percebemos, eu particularmente, que o jornal estava um pouco

---

<sup>170</sup> O *Akademia* foi descontinuado por falta de financiamento, já que fazia parte de um projeto do Labcom financiado com recursos da Fundação de Ciência e Tecnologia (FCT). O assunto tornou-se, inclusive, um tema para enquete ou inquérito, no qual era solicitado aos usuários que dessem sua opinião sobre o trabalho realizado. Tanto o professor António Fidalgo, bem como outros membros da equipe do *Akademia*, a exemplo de João Simão, declararam que o novo modelo colocado pelo produto teve pouco tempo para fidelizar uma comunidade de usuários, o que, de certa forma, justificava a baixa participação. No entanto, segundo Simão, já perto do fechamento, a participação e o interesse dos usuários começava a crescer, tanto pelo modelo do produto como pelo tema. Pudemos observar e analisar o material do *Akademia*, assim como realizar entrevistas com a equipe idealizadora e gestora do produto, *in loco*, durante o período do estágio-sanduíche, realizado de março a dezembro de 2005, no Labcom, sob a supervisão do professor António Fidalgo. Através do site [www.labcom.ubi.pt](http://www.labcom.ubi.pt) é ainda possível navegar pelos arquivos do *Akademia*.

<sup>171</sup> [www.urbi.ubi.pt](http://www.urbi.ubi.pt).

<sup>172</sup> As declarações de João Canavilhas foram concedidas à autora em entrevista realizada no dia 28 de Abril de 2006, quando tivemos a oportunidade de estar mais uma vez no Labcom/UBI, acompanhando a etapa final da reformulação do *Urbi et Orbi*.

distante daquilo que, conceitualmente, estávamos trabalhando”. Entre as principais vantagens da mudança realizada, Canavilhas cita que, com o uso de bases de dados, puderam-se criar as “Tops” (matérias mais lidas), “Jornal do cidadão” (para agregar o material produzido pelo conjunto dos usuários), e também as “Últimas”. “Assim, passamos a contar com três níveis de periodicidade: o imediato (“Últimas”); o semanal (edição *on-line*, com atualização todas as terças); e o mensal (edição impressa de todo o conteúdo)”.

Outras melhorias foram a incorporação do áudio e do vídeo, da técnica do *podcasting*<sup>173</sup>, de infográficos e, claro, automatização dos processos de edição e publicação, assim como o dinamismo conferido ao *site*. Para Eduardo Alves, chefe de redação, o processo de edição se tornou mais flexível, favorecendo o trabalho simultâneo. Com isso, o processo não afunila mais na mão de uma pessoa, pois agora repórteres, editores, diretores, dispõem da mesma ferramenta, que permite escrever e publicar e a página é gerada automaticamente<sup>174</sup>.

“Enfim, o novo *Urbi et Orbi* é a experimentação, materialização da teoria. Não basta ter um modelo a rodar”, assegura João Canavilhas. Ainda abordaremos as mudanças efetuadas neste produto, bem como trataremos dos outros *sites* que integram o conjunto de casos trabalhados na tese, neste e no próximo capítulo.

---

<sup>173</sup> De acordo com Canavilhas, o *Urbi et Orbi* foi o primeiro jornal semanal *on-line* português a disponibilizar para os seus leitores sons, vídeos e imagens em *podcasting*.

<sup>174</sup> Entrevista concedida à autora em 18/05/2006.



Figura 23 – Reformulação permitiu ao *Urbi* deixar de ser estático para se tornar um produto dinâmico. Jun.2006.

O pesquisador brasileiro, Elias Machado (2004a, 2004b, 2004c, 2006), ao refletir sobre o potencial das bases de dados no jornalismo digital, vai afirmar que elas constituem um formato para esta modalidade. É em consonância com o princípio da transcodificação (segundo o qual todos os objetos da nova mídia podem ser traduzidos para outros formatos), descrito por Lev Manovich (2001), que ele defende a hipótese dos bancos de dados como forma cultural com estatuto próprio no jornalismo digital. Para o autor (2004a, p.02), as BDs desempenham três funções simultâneas e complementares: a) de formato para a estruturação da informação; b) de suporte para modelos de narrativa multimídia; e c) de memória dos conteúdos publicados, o que o leva a considerar os bancos de dados como um formato no jornalismo digital.

Machado (2004a) argumenta que, de igual modo à narrativa literária ou cinematográfica e a um plano arquitetônico na Modernidade, o banco de dados emerge como a forma cultural típica para estruturar as informações sobre o mundo/realidade na cultura dos computadores. Segundo ele:

A plena incorporação pelas organizações jornalísticas da lógica dos bancos de dados depende da utilização casada das funções de modelo de estruturação da informação, espaço para criação de

narrativas e lugar para a ativação da memória (...) Para que o princípio da transcodificação seja aplicável ao jornalismo digital, o banco de dados deve servir tanto como um espaço para a experimentação de formas diferenciadas de narrativa multimídia, quanto como uma fonte de atualização do presente vivido à luz da memória armazenada (MACHADO, 2004a, p.12-13).

E o primeiro passo para constituir uma estética própria para as organizações jornalísticas, de acordo com o pesquisador, passa, justamente, pela compreensão que nas novas mídias os elementos constitutivos da narrativa são formatados como bancos de dados. Além da compreensão por parte das próprias organizações, acrescentamos que também os profissionais jornalistas devem ampliar o seu entendimento e capacidade cognitiva sobre o potencial dos bancos de dados, pois o desempenho das suas funções certamente irá requerer tal conhecimento. Afinal, como anuncia Elias Machado, o futuro das organizações jornalísticas nas redes permanece condicionado à capacidade de traduzir as habilidades potencializadas pelos bancos de dados para automaticamente armazenar, classificar, indexar, conectar, buscar e recuperar vastas quantidades de dados em tipos criativos de narrar, por exemplo, o passado como se fosse um presente projetado em direção ao futuro.

Estamos de acordo com a proposição de Elias Machado (2004a, 2004b, 2004c) e, em artigos publicados naquele mesmo ano (BARBOSA, 2004a, 2004b, 2004c), conjecturamos, também a partir das formulações de Manovich (2001) e, de certo modo, do próprio Machado, que o emprego das bases de dados no jornalismo digital tinha o potencial de gerar uma nova metáfora, capaz de trazer nova luz no sentido de se superar a metáfora do jornal impresso que, desde os primeiros anos de experimentação do jornalismo no suporte digital até agora, ainda é empregada pelos mais diferentes tipos de *sites* noticiosos.

As investigações de Elias Machado sobre o uso das BDs resultaram, inclusive, na realização de uma comunicação coordenada sobre o tema no II Encontro Nacional de

Pesquisadores em Jornalismo da Associação Brasileira de Pesquisadores em Jornalismo (SBPJor), em 2004. Dela participaram António Fidalgo (Universidade da Beira Interior, Portugal), o próprio Elias Machado (atualmente, professor na Universidade Federal de Santa Catarina), Claudia Quadros (Universidade Tuiuti do Paraná), Walter T. Lima Júnior (Universidade Metodista de São Paulo) e a autora desta tese. Os desdobramentos daquele encontro oportuno só vieram confirmar a pujança de tal temática que continua a ser explorada pelos participantes, nas pesquisas que vêm desenvolvendo desde então ou motivando outros pesquisadores brasileiros, além do que se mostram afinados com investigações de outros estudiosos, na Espanha, no Chile, nos Estados Unidos, entre outros países.

No livro lançado recentemente, que reúne, entre outros, os artigos escritos entre 2004 e 2005, Elias Machado (2006) discute algumas das particularidades das bases de dados para o jornalismo digital, e as relações existentes entre sistemas de publicação e modelos de narrativa, entre arquitetura da informação e narrativa, os elementos que constituem o roteiro da narrativa multimídia, e as funções exercidas pelos diversos atores sociais envolvidos em processos de produção de conteúdos jornalísticos. Uma das diferenças fundamentais para os produtos jornalísticos digitais estruturados em bases de dados, segundo o autor, é que:

(...) todas as etapas do processo de produção são articuladas em torno das possibilidades disponibilizadas pelas bases de dados. Ao tornar o produto jornalístico digital e as relações de produção muito mais racionalizadas e flexíveis, as Bases de Dados permitem a reestruturação destas relações no jornalismo, podendo transformar por completo as funções tanto dos jornalistas, quanto dos usuários destas publicações (MACHADO, 2006, p.08).

Quanto às consequências da incorporação das BDs no processo de apuração, na circulação e na comercialização das informações, Elias Machado diz tratar-se de um dos

desafios para trabalhos futuros. Como é um tema bastante relevante, certamente, estudos específicos subseqüentes deverão evidenciar os resultados dessa incorporação para o jornalismo.

Continuando com o exame do estado da arte sobre o emprego de bases de dados no jornalismo, passaremos à perspectiva trabalhada por Claudia Quadros (2004, 2005). Para a investigadora, a maneira como essa tecnologia vem sendo aplicada ressalta o poder que elas possuem para gerar conhecimento e abrir novas possibilidades para diversos setores da sociedade. Sua preocupação está em apontar como os usuários podem interagir com a mídia por meio da utilização das BDs por parte das empresas jornalísticas. Neste sentido, ela afirma: “As bases de dados de linguagens verbais, visuais (fotografias, vídeos, gráficos, infográficos, multimídia, etc.) e sonoras podem ser para o jornalismo uma espécie de representação da democracia ao seu usuário” (QUADROS, 2005, p.05). Segundo a autora, o jornalismo pode garantir, com uma base de dados dinâmica, conteúdos e produtos de qualidade aos usuários de diversos meios. Pois, uma vez que as informações estejam digitalizadas e armazenadas, elas podem ser recuperadas a qualquer momento, resgatando a memória a serviço do público.

Sendo assim, ela esclarece: além de melhorar a qualidade dos conteúdos, a base de dados dinâmica permite disponibilizar novas formas de interação com o público e criar outros produtos e serviços. Por isso, adverte Quadros (2005, p.06), as empresas de comunicação devem saber que esse tipo de armazenamento de conteúdos pode ser utilizado não apenas como instrumento pelos mais diversos veículos, mas também para a criação de novos produtos digitais mais criativos e estabelecer uma relação mais estreita com o usuário, que encontrará inúmeras possibilidades nessa imersão digital.

Em qualquer desses casos, as implicações nas rotinas produtivas e de ordem editorial são facilmente percebidas. A revista impressa

pode, por meio de sua base de dados, criar alguns produtos, como, por exemplo, um CD – um meio digital – sobre a cobertura completa das quatro últimas olimpíadas e, assim, atender às necessidades do leitor interessado em resgatar uma memória específica. Esse tipo de promoção, cada vez mais comum no mercado editorial, é uma forma de aliar conteúdo e negócio sem comprometer a ética. A base de dados dinâmica no jornalismo digital contribui na exploração das potencialidades da hipermídia e garante a interatividade do usuário a ponto de esse assumir os papéis de receptor e emissor de uma mensagem (QUADROS, 2005, p.06).

Já Walter T. Lima Júnior (2004, 2005, 2006) tem se dedicado a estudar a técnica do *data mining* em como ela pode auxiliar o jornalismo na melhoria da qualidade da informação pesquisada em bancos de dados e na obtenção de relações ‘invisíveis’ de temas e contextos. Mais recentemente, tem levado a cabo seu projeto de construção de uma base de dados para pesquisa e validação qualitativa de fontes de informação jornalística (2006). Este projeto integra a pesquisa de pós-doutorado, intitulada “Formação das bases conceituais e criação de UML (Unified Modeling Language) visando à produção de *software* para pesquisa e validação qualitativa de fontes de informação jornalística”<sup>175</sup>.

Sobre a aplicação da mineração de dados no jornalismo para extrair padrões válidos – por exemplo, se o índice de desemprego diminui quando se aproxima uma eleição e porque isso acontece –, Lima Júnior (2004, p.10) entende que é preciso que os bancos de dados sejam precisos e não históricos e que tenham uma certa inteligência artificial para lidar com as modificações semânticas das palavras, entre outras.

---

<sup>175</sup> O pós-doutorado está sendo realizado na Universidade Metodista de São Paulo (UMESP). No 4º Encontro Nacional dos Pesquisadores em Jornalismo (SBPJor), realizado de 05 a 07 de Novembro de 2006, em Porto Alegre, Lima Júnior apresentou a primeira fase da pesquisa. Esta fase tem como objetivo “simplificar as formas de solucionar problemas de decisão na escolha de fontes jornalísticas, conceituando variáveis e as transferindo para uma versão computacional, constituídas de sistemas inteligentes digitais” (LIMA JÚNIOR, 2006, p. 16).

A precisão requerida para eliminar distorções e inconsistências nas bases de dados é uma condição fundamental, assim como também o é a habilidade requerida dos profissionais para executar essa tarefa. Para Lima Júnior, eles terão que possuir ‘conhecimento de mundo’; que as máquinas ainda não dispõem.

No jornalismo, assinala, as *databases* (fontes de dados) seriam compostas por bancos de dados com matérias publicadas (históricos), listas de conteúdo ou resumos de CD e DVD’s e bancos de dados disponíveis em redes (internet ou intranet), mas que tivessem consistência nas informações disponíveis (dados precisos e pertinentes), remoções de ruídos e redundância. “Além disso, também têm que ser mais amplos, ou seja, deixando de serem apenas repositórios de textos e fotos. Poderiam conter vídeo (por palavras-chave controladas, resumos, dados sobre sonoras, *offs* e videografia) e áudio (palavras-chave controladas, resumos, dados sobre sonoras e *offs*)”, afirma Walter Lima Júnior (2004, p.12).

Outra maneira de empregar as técnicas do *data mining* se refere aos modos diferenciados para a apresentação/visualização dos conteúdos. Como mencionamos no primeiro capítulo, há produtos como o *Newsmap*<sup>176</sup>, o *News Explorer*<sup>177</sup>, o *News Brief*<sup>178</sup> (*European Media Monitor*), entre outros ainda funcionando como versões beta (teste) – a exemplo do *Newzbubble*<sup>179</sup> –, que já aplicam o *data mining* com a intenção de promover visualizações menos convencionais para as informações jornalísticas. O que seria, desta maneira, uma aplicação da estética base de dados ou da *database aesthetics* no jornalismo digital.

---

<sup>176</sup> [www.newsmap.com](http://www.newsmap.com).

<sup>177</sup> <http://press.jrc.it/NewsExplorer/dayedition/en/latest/html>.

<sup>178</sup> <http://press.jrc.it/NewsBrief/worldedition/en/en.html>.

<sup>179</sup> <http://www.newzbubble.com/index>.

### 3.1.2.1 O valor da informação estruturada e a automatização dos processos

Adrian Holovaty (2005, 2006a, 2006b), jornalista, programador e editor de inovações editoriais do *Washingtonpost.com*<sup>180</sup>, tem sido um dos pesquisadores atentos ao valor da informação estruturada para o jornalismo. Embora, por outro lado, perceba a existência de algumas dificuldades para a sua implementação em decorrência desse campo ser “implacavelmente desestruturado”. Segundo afirma Adrian Holovaty, um dos problemas está exatamente naquela que é a matéria-prima do trabalho jornalístico: “The primary product of journalists - the *news story* - is just a giant blob of text” (HOLOVATY, 2006a, p.01)<sup>181</sup>.

Conforme explica, uma notícia não pode ser dividida facilmente em pedaços consistentes, pois não há fatos em lugares previsíveis. Deste modo, os conjuntos de notícias, mesmo quando publicadas num mesmo jornal ou em outra publicação e escritas pelo mesmo repórter, não são consistentes nem legíveis para um computador moldá-las, adaptá-las. Adrian Holovaty explica que ainda que as notícias possam ter metadados<sup>182</sup> (destaques, datas de publicação, nomes dos autores das matérias, categorização), a essência das mesmas é o texto e ele não é estruturado.

---

<sup>180</sup> [www.washingtonpost.com](http://www.washingtonpost.com) .

<sup>181</sup> “O produto primário dos jornalistas – as notícias – é apenas um borrão gigantesco de texto”. (t.A.).

<sup>182</sup> São dados acerca dos dados ou informação acerca da informação, ou seja, contêm informação semântica sobre os dados. Raymond Colle (2002, p.34) os define como: dados sobre outros dados, ferramentas que guiam os usuários aos dados tanto para encontrar informação pontual como para extrair informação sobre o conjunto e que provêm um contexto que pode ser de grande importância para uma melhor interpretação das informações. As co-relações entre as notícias inseridas numa base de dados, considerando a classificação por meio de diversos campos, bem como as possibilidades combinatórias entre elas permitirá produzir, extrair novas informações, novos dados, que vão gerar mais conhecimento, mais contexto, sobre os eventos, segundo a nossa perspectiva.

A ausência de estrutura, por outro lado, certamente tem algum sentido, pois, como argumenta o autor, uma notícia é planejada para ser “consumida” por humanos, não por computadores. Ainda assim, Holovaty observa que boa parte das informações sobre crimes, obituários, dentre outras, segue um certo padrão e, desta maneira, podem ser ricamente exploradas por bases de dados.

Para Holovaty (2005, 2006a, 2006b), a combinação de BDs com linguagens de programação pode ajudar a introduzir algum valor estruturado para as notícias e para os conteúdos jornalísticos de um modo geral. A *eXtensible Markup Language (XML)*<sup>183</sup>, associada às bases de dados, por exemplo, confere dinamismo e automatização ao processo de produção jornalístico. Ele tem defendido que o jornalismo, e os jornalistas, precisam tirar partido dos computadores, das BDs e das linguagens de programação para a escrita, edição e a apresentação das informações jornalísticas. Neste sentido, o autor cita e explica como três tarefas básicas no jornalismo podem ser automatizadas<sup>184</sup>. São elas:

- Obtenção/Coleta de informação (*Gathering information*). Ele comenta que esse trabalho que envolve falar com fontes, examinar documentos, fazer fotografias, etc., para compor as peças informativas, pode ser realizado conforme este exemplo implementado por ele no *Washingtonpost.com*: um programa foi desenvolvido para checar, várias vezes ao dia, os *web sites* do governo americano para encontrar

---

<sup>183</sup> Uma das principais características desta linguagem é permitir desvincular a forma do conteúdo. Segundo Arthur Afonso Sousa (2002, p.07), a *XML* é uma (meta) linguagem de marcação de documentos completamente independente das plataformas *hardware* e *software* que a utilizam, sendo um padrão aberto. Conforme Jim Hall (2001, p.74) e Stephen Quinn (2002, p.133), a *XML* é uma linguagem de estruturação de dados que habilita a produção de documentos no qual o dado é separado da forma ou da apresentação dele. A *XML* é a linguagem que torna possível, por exemplo, a troca de dados entre aplicações heterogêneas, fundamentalmente na internet, pois foi projetada para os requerimentos da WWW. Por isso, ela é apontada como novo *standard* para a representação e a permuta de dados na rede tendo sido proposta pelos pesquisadores do *W3C – World Wide Web Consortium*.

<sup>184</sup> Essas tarefas foram descritas por ele na entrevista à Robert Niles da *Online Journalism Review (OJR)*, em 05/06/2006.

informações sobre votações de projetos, leis, etc. Quando encontra alguma, o programa obtém o dado e a salva em uma BD. Programa similar também funciona no *site* [www.chicagocrime.org](http://www.chicagocrime.org)<sup>185</sup>, criado por Holovaty. Neste caso, o programa “varre” o *web site* do Departamento de Polícia de Chicago para obter todas as informações sobre os crimes ocorridos na cidade;

- Edição/tratamento da informação (*Distilling information*). Do mesmo modo que um editor pode fazer o julgamento editorial para decidir quais fatos em um acontecimento são mais importantes, um jornalista-programador<sup>186</sup> decide quais *queries* (perguntas) deveriam ser feitas. No *Chicagocrime.org*, diz, foi ele quem decidiu se seria útil para os usuários navegar pelos tipos de crime, código postal e distrito da cidade. Já na base de dados do *Washingtonpost.com* sobre as votações, foi decidido que seria mais útil navegar pela lista das votações que aconteciam no final do dia de trabalho no congresso e pela lista de membros do congresso que estavam faltando à maioria das sessões;

- Apresentação da informação (*Presenting information*). Segundo Holovaty, esta é a etapa particularmente mais complexa de ser automatizada, porque implica a criação de *web sites* – ou, diríamos, interfaces para exibir os conteúdos. Por isso, ele adverte, é necessário considerar todas as possíveis permutações, trocas, dos dados. No caso do *Chicagocrime*, ele precisou levar em conta a perda de dados e como resolver a questão da exibição de informações que poderiam não estar disponíveis à época do crime.

---

<sup>185</sup> Este *site* ganhou, em 2005, o *Knight-Batten Award Innovations in Journalism* (<http://www.jlab.org/batten05winners.shtml>), concedido pelo The Institute for Interactive Journalism da University of Maryland (US) e foi considerado pelo *The New York Times* uma das melhores idéias de 2005. Adrian Holovaty possui outros projetos para sites automatizados de informação jornalística e já trabalhou em outros produtos digitais, entre os quais, o [www.lawrence.com](http://www.lawrence.com), *site* de entretenimento local de Lawrence (Kansas), o qual ele considera bastante inovador. Tanto o *Chicagocrime*, como o *site* do *Lawrence*, entre outras aplicações desenvolvidas para o *Washingtonpost.com*, foram criadas por Holovaty com o *Django* (<http://www.djangoproject.com/>), um *software* livre baseado na linguagem Python e que foi desenhado para ser implementado em redações jornalísticas.

<sup>186</sup> Ele não gosta do termo e diz que realmente se precisa de um nome melhor.

É por conta de situações como esta que o jornalista e programador norte-americano relativiza e assinala que, mesmo sendo possível construir *sites* totalmente automatizados, outros produtos poderão funcionar com apenas uma parte do processo jornalístico automatizado. Um exemplo é o especial *Faces of the Fallen*, do *Washingtonpost.com*. O *site*, que lista todos os soldados norte-americanos que têm morrido no Iraque (e no Afeganistão) desde o início da ocupação norte-americana, teve todo o processo de coleta e verificação das informações feito pela equipe de repórteres da redação do jornal.

Por outro lado, a apresentação do conteúdo (que pode ser navegado segundo uma variedade de parâmetros, como idade, data de nascimento e de morte, cidade em que moravam os soldados e qual a sua patente militar, entre outros) foi feita de modo totalmente automatizado). "Doing journalism through 'computer programming' is just a different way of accomplishing these goals. Namely, the technique favors automation wherever possible"<sup>187</sup>, ratifica Holovaty.

The screenshot shows the 'Faces of the Fallen' section on the Washington Post website. At the top, there's a navigation bar with various categories like News, Politics, Opinions, Sports, Arts & Living, Discussions, Photos & Video, City Guide, Classifieds, Jobs, Cars, Real Estate, and Shopping Deals. The main heading is 'Faces of the Fallen' with a sub-heading 'U.S. service members who died in Operation Iraqi Freedom and Operation Enduring Freedom'. Below the heading, there's a search bar and a filter for 'View Service Members by: Age, Year of Death, Home State, Military Branch, Advanced Search'. The main content area displays a grid of 100 small portraits of fallen soldiers, organized by date of death. A sidebar on the right shows 'Total Fatalities' as 3,141, with a breakdown for Operation Iraqi Freedom (2,801) and Operation Enduring Freedom (340). Below this, there's a 'Related Coverage' section with several article links.

Figura 24 - *Faces of the Fallen* do *Washingtonpost.com*: automatização para a apresentação do conteúdo. Nov.2006.

<sup>187</sup> "Fazer jornalismo através da programação de computador é apenas um modo diferente de cumprir essas metas. Nomeadamente, a técnica torna qualquer automação possível". (t.A.).

A seguir, a imagem mostra a interface criada para a base de dados sobre as votações no congresso norte-americano. O *site* recebeu o *Knight-Batten Awards 2006 for Innovations in Journalism*<sup>188</sup>, da University of Maryland (US). A BD também está integrada no processo de produção do *Washingtonpost.com*.

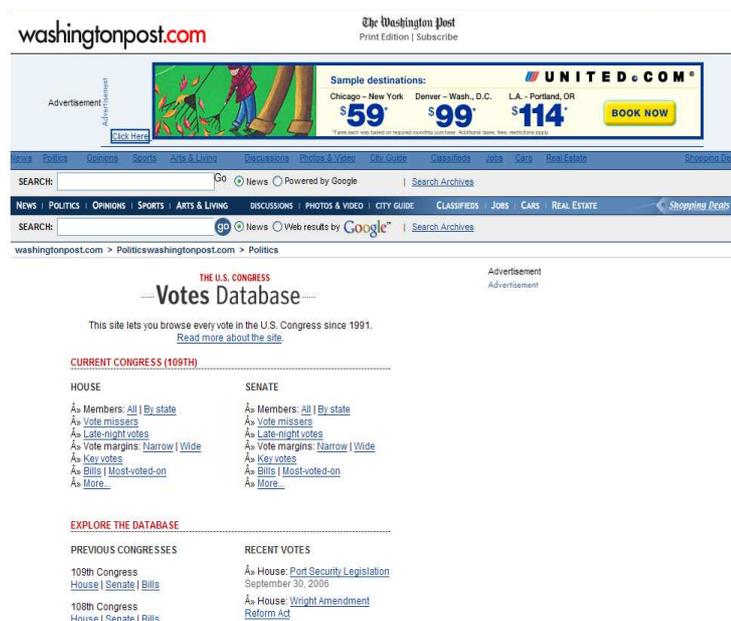


Figura 25 - *Votes Database*, também do *Washingtonpost.com*, agrega todas as informações sobre votações ocorridas no congresso dos EUA. Nov.2006.

Ainda sobre o aspecto da automatização de produtos jornalísticos digitais, cabe referenciar que um dos maiores exemplos de *site* funcionando totalmente de modo automatizado é o *Google News*, agregador de notícias, surgido em setembro de 2002. Ele opera com algoritmos que “varrem”, isto é, vasculham 4.500 publicações noticiosas em toda a web e, de acordo com a metodologia empregada, organiza/indexa as informações, distribuindo-as em categorias e subordinando os assuntos que estão sendo

<sup>188</sup> <http://www.j-lab.org/ba06finalists.shtml>.

mais citados ou acessados, com atualização de 15 em 15 minutos<sup>189</sup>. Há versões do *Google News* em vários idiomas, entre eles o português<sup>190</sup>, que ainda aparece com edição beta (teste) desde o seu lançamento em novembro de 2005, listando, por enquanto, apenas 200 fontes, incluindo tanto aquelas provenientes de publicações brasileiras, como aquelas de Portugal.

Sem a intervenção de repórteres e editores, o processo de seleção e julgamento jornalístico sobre a relevância das fontes referenciadas fica a cargo dos algoritmos, também chamados de *crawlers Googlebot*: os robôs do *Google* que buscam informações nos *sites*. Esta metodologia tem ocasionado alguns problemas quanto à credibilidade e legitimidade das fontes relacionadas e, por outro lado, rendido críticas ao modelo *Google News*<sup>191</sup>. Para atribuir relevância às notícias que indexam nas seções “Top Stories”, “World”, “US”, “Business”, “Sci/Tech”, “Sports”, “Entertainment”, “Health”, “Most popular”, os algoritmos do *Google* consideram fatores tais como o maior número de matérias a respeito deles e o quanto são repetidas as referências. Trata-

---

<sup>189</sup> O *Google* é um caso paradigmático neste aspecto, por ser industrial, competitivo e ter provado sua capacidade comercial. No entanto, há outros projetos similares desenvolvidos no âmbito acadêmico, a exemplo do *Newsblaster* ([www.newsblaster.com](http://www.newsblaster.com)), criado a partir de técnicas de inteligência artificial no Departamento de Ciências da Computação da Universidade de Colúmbia, EUA. Ele coleta informações em várias fontes da web, examina as notícias publicadas, seleciona algumas e as sintetiza, usando técnicas de processamento de linguagem natural, associadas a bases de dados, para redigir e publicar resumos jornalísticos. Outro exemplo é o *News in Essence* (<http://lada.si.umich.edu:8080/clair/nie1/nie.cgi>), da Universidade de Michigan, EUA, que coleta e resume notícias de fontes diversas sobre tópicos específicos (LAGE, 2002).

<sup>190</sup> [http://news.google.com/nwshp?ned=pt-BR\\_br](http://news.google.com/nwshp?ned=pt-BR_br).

<sup>191</sup> Sobre o assunto ver Staci D. Kramer (2002) e Isabel Aguilera (2006), por exemplo. Em 18/09/2006, uma corte de Bruxelas, a partir de ação movida por alguns diários belgas, obrigou o *Google* a remover o conteúdo de todos os jornais publicados em francês e alemão na Bélgica do seu serviço *Google News*, que estaria lucrando sem pagar nada para usar tais conteúdos (MELLER, 2006). Os jornais brasileiros membros da Associação Nacional de Jornais (ANJ), no seu 6º congresso realizado no final de Agosto de 2006, em São Paulo, decidiram que passariam a cobrar de sites como o *Google News* uma parte dos lucros. A notícia, conforme o *Jornalistas da Web* ([www.jornalistasdawe.com.br](http://www.jornalistasdawe.com.br)) divulgou em 04/09/2006, repercutiu em sites estrangeiros, como o britânico [www.dotJournalism](http://www.dotJournalism). Em Setembro de 2006, o *Google News* fechou acordo com a *Associated Press* para usar as informações produzidas pela agência de notícias. Por outro lado, a *France Press* também move ação contra o *Google* por uso indevido de suas informações.

se do sistema *PageRank* (que atribui uma pontuação), e significa que, quanto mais *links* existirem apontando para uma página, maior será seu grau de importância no *Google*.

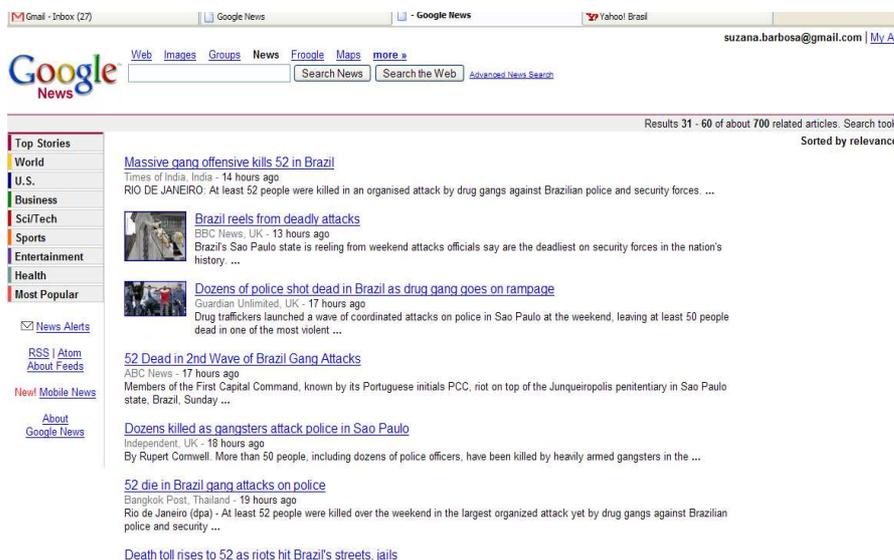


Figura 26 – *Google News* é um dos mais conhecidos *sites* agregadores de notícias, funcionando de modo completamente automatizado. Maio.2006.

Entre os produtos que formam o conjunto de casos que integram o *corpus* empírico trabalhado nesta tese, estão outros *sites* que funcionam de modo inteiramente automatizado, com algoritmos e tecnologias específicas, técnicas de inteligência artificial, entre outras, associadas às bases de dados. Em geral, são produtos agregadores de notícias<sup>192</sup>, ou que também produzem sínteses, resumos, a partir da indexação de informações produzidas por outros *sites* jornalísticos, também remetendo às fontes originais para a leitura dos textos na íntegra. São eles: *Newsmap*, *Newzbubble*, *News Explorer* e *News Brief*, *Topix.Net*, e também a *Notibits*.

<sup>192</sup> Há, contudo, agregadores de notícias que operam com jornalistas para fazer a seleção da informação que publica. Um exemplo é o *DailySource* (<http://dailysource.org/>), criado em 2005 e ainda funcionando em versão beta.

The screenshot shows the Topix.net website interface. At the top, there is a navigation bar with categories: News Front Page, Forums, Blogs, Local, US, World, Business, Sports, Entertainment, Offbeat, Classifieds, and Browse All Topix. Below the navigation bar is the Topix.net logo and a search bar. The main content area is titled "News Front Page" and displays a "TOP STORY" about car bombs in Baghdad. Below this, there are sections for "LATEST NEWS" with various headlines, "Recent Comments" with a list of user comments and their counts, and "Ads by Google" featuring travel-related advertisements for Rio de Janeiro.

Figura 27 - O Topix.net, criado em 2002, tem tecnologia própria para oferecer conteúdos locais, o foco da sua oferta informativa. Dez.2006.

Sobre o *Newsmap*, é importante salientar que ele trabalha exibindo tópicos de notícias do *Google News*, em uma forma visual específica e dentro de um espaço limitado<sup>193</sup>. O usuário pode rapidamente identificar quais notícias estão sendo publicadas com maior cobertura, vendo o mapa da região, país, categoria temática ou tempo. O objetivo do *Newsmap*, criado em 2004, é o de ser um serviço complementar, ajudando os usuários a analisar e a comparar a cobertura feita sobre os acontecimentos pelos mais diversos *sites* jornalísticos, seja nos Estados Unidos, Canadá, Austrália, Áustria, França, Espanha, Itália, Nova Zelândia ou Reino Unido.

<sup>193</sup> O idealizador do *Newsmap*, Marcos Weskamp, tomou como inspiração para o formato de visualização das informações o modelo *tree-map*, desenvolvido por Ben Shneiderman, em 1999. Shneiderman, fundador do *Human Computer Interaction Laboratory*, da University of Maryland (US), é professor daquela instituição e um dos precursores na investigação sobre visualização da informação. Trata-se de um campo de pesquisa distintivo, que tem se expandido rapidamente e se tornado um campo interdisciplinar de pesquisa. Conforme escreve Shneiderman no prefácio do livro *Information Visualization: Beyond the Horizon*, de Chaomei Chen (2004), a visualização da informação representa uma tendência estabelecida no moderno *design* de interface para o usuário (SHNEIDERMAN, 2004, p.01).



brasileiro *Folha Online* integram o *ranking* de notícias em seus respectivos *sites*, trabalhando a visualização desses dados de modos diversos.

The image shows a screenshot of the Folha Online website. At the top, there's a banner for 'FOLHA ONLINE' with the website address 'www.folha.com.br'. Below that, it says 'FOLHA DE S. PAULO' and 'Quinta-feira, 28 de dezembro de 2006'. The main navigation bar includes links like 'Sobre o site', 'Fale com a gente', 'Assine a Folha', 'Atendimento ao Assinante', and 'Anuncie'. The main content area features a headline: 'Mais dois ônibus são atacados; onda de violência segue no Rio'. A red arrow points to a sub-headline: 'TAM nega fretamento e aumenta número de funcionários para amenizar crise'. To the right, there's a search bar and a 'Lidos' section with a list of 5 items. At the bottom, there's a 'folhashop' advertisement.

Figura 29 – *Folha Online* é um dos poucos *sites* nacionais que exibem na *home* os “+ Lidos”. Dez.2006.

Ressaltamos, ainda, que o *Portal Terra* e o *blog Deu no Jornal* da *ONG Transparência Brasil*, também trabalham conteúdos gerados automaticamente a partir de informações contidas em bases de dados. No *Terra*, há o subcanal *Fatos Históricos* (está inserido no canal “Notícias”<sup>195</sup>), apresentado com o título “O que fez esse dia entrar para a história”. O subcanal lista os eventos-chave ocorridos ao longo dos séculos, organizando-os por meses, podendo, em alguns casos, indexar acontecimentos desde o século XII até o XXI e, destacando no alto da página, um evento determinado<sup>196</sup>.

<sup>195</sup> <http://noticias.terra.com.br>.

<sup>196</sup> Entre 2003 e 2004, o *Estadão* ([www.estadao.com.br](http://www.estadao.com.br)) fez experiência com o material de arquivo num canal exclusivo chamado “Diário do Passado”, demonstrando uma das possibilidades para se trabalhar conteúdos de edições passadas, a partir de combinações e cruzamentos entre informações jornalísticas inseridas na sua base de dados.

**notícias**

Notícias > Fatos Históricos

notícias por e-mail fale conosco rss

**ESTACAÇÃO VINHO**  
O melhor do Vinho na Internet

**punto final**  
Malbec Varietal 2003  
87 pts Wine Spectator

**Fatos Históricos**

**Fatos históricos do dia 31 de outubro**

**Acidente mata 99 em SP**  
Em 31 de outubro de 1996, a queda de um avião da companhia aérea TAM, em São Paulo, matou 99 pessoas, sendo 96 que estavam no avião e três que estavam em terra. O acidente ocorreu por um problema no reverso da turbina. A aeronave caiu em cima das casas no bairro de Jabaquara, após decolar.

Outubro						
1	2	3	4	5	6	7
8	9	10	11	12	13	14
15	16	17	18	19	20	21
22	23	24	25	26	27	28
29	30	31				

**Outros Meses**

- » Janeiro
- » Fevereiro
- » Março
- » Abril
- » Maio
- » Junho
- » Julho
- » Agosto
- » Setembro
- » Outubro
- » Novembro
- » Dezembro

**1517** - Martinho Lutero prega suas 95 teses na porta da Igreja de Wittenberg, marcando o começo da Reforma Protestante na Alemanha.

**1714** - Jorge I da Inglaterra é coroado rei da Grã-Bretanha e da Irlanda.

**1790** - Surge o *Penódico de Havana*, primeiro jornal literário-econômico publicado na capital cubana.

**1794** - O físico e químico inglês John Dalton descobre o daltonismo, anomalia ótica que impede a pessoa de distinguir certas cores, principalmente o vermelho e o verde.

**1943** - Em Londres, diante do prédio da Aliança Francesa, o general De Gaulle homenageia os escritores franceses resistentes.

**1952** - A primeira bomba de hidrogênio é detonada pelos Estados Unidos nas ilhas Marshall, no atol Eniwetok, região do Pacífico.

**1954** - A Frente para a Libertação Nacional da Argélia começa a revolução contra a dominação francesa. A independência só seria conquistada em 1962.

**1969** - O presidente norte-americano Lyndon Johnson anuncia um completo cessar-fogo de bombardeios contra o Vietnã do Norte.

**1972** - Polícia paulista desvenda ramo brasileiro da máfia italiana e prende o chefe Tomaso Buscetta.

**1974** - Grã-Bretanha, França e Estados Unidos vetam uma moção das Nações Unidas de expulsar a África do Sul.

**1978** - Aprovada a nova constituição espanhola.

**1979** - O jornalista Fernando Gabeira lança o livro *O que é isso, companheiro?*, que se tornou filme e inspirou o *Casa de Melhor Filme Estacação*.

Figura 30 – *Portal Terra* – Notícias: o subcanal “Fatos Históricos”. Nov.2006.

O *blog Deu no Jornal* da ONG Transparência Brasil, por seu turno, é uma base de dados que reúne mais de 90 mil matérias sobre a temática “Corrupção” publicadas em 63 jornais e revistas de todo o Brasil. A partir dessas fontes, o editor Marcelo Soares faz uma análise daquelas notícias consideradas mais relevantes e as publica. O *blog*, criado em 2004, disponibilizou em abril de 2006 a opção de se construir mapas de relacionamento entre as pessoas e os assuntos mencionados nas matérias acumuladas em sua BD, gerando outras possibilidades de visualização para os dados.

**DEU NO JORNAL**

BLOG COM A ANÁLISE DAS PRINCIPAIS NOTÍCIAS SOBRE CORRUPÇÃO PUBLICADAS EM 63 JORNAIS E REVISTAS DE TODO O BRASIL, REUNIDAS NO BANCO DE DADOS DEU NO JORNAL, DA TRANSPARÊNCIA BRASIL. EDITADO POR MARCELO SOARES, COORDENADOR DO PROJETO.

Segunda-feira, 11 de Dezembro de 2006

**Pinochet e o direito de acesso**

O general Augusto Pinochet, ditador do Chile por quase 20 anos após o 11 de setembro de 1973, é um personagem importante na história do direito de acesso a informações públicas. Pinochet morreu sem ser condenado, mas o processo contra ele trouxe benefícios ao nosso conhecimento histórico.

O escritor Fernando Morais escreveu para a revista *Playboy*, em 1999, um texto definitivo sobre o evento que desencadeou a grande contribuição involuntária de Pinochet para o direito de acesso. É um perfil do juiz espanhol Baltazar Garzón, o homem que mandou prender o ditador. Republicado no livro "Cem Quilos de Ouro", ele começa assim:

*Eram onze da noite de 16 de outubro de 1998, e o general chileno Augusto Pinochet, 82 anos, cochilava num quarto do oitavo andar da London Clinic, no número 20 da Devonshire Place, no centro da capital britânica. Não entendeu o que estava acontecendo quando, despertado pelo barulho, viu um homem de cabelos curtos e terno azul-marinho de pé, ao lado de sua cama, com uma folha de papel na mão. Era a primeira vez em muitos anos - uns trinta, quem sabe? -*

**EXCELENCIAS**  
Como se comportam os nossos deputados.  
TransparênciaBrasil

**CATEGORIAS**  
Acesso a Informações Públicas  
Eleições  
O melhor do Jornalismo  
Pra relaxar  
Dicas de uso

**TRANSPARÊNCIA**  
Deu no Jornal  
Transparência Brasil  
As Claras  
Multibusa  
Licitações de SC

**FERRAMENTAS**  
Políticos do Brasil  
Teia Eleitoral  
TSE - Divulgação de Candidaturas  
Doxa - Juperj  
Contas de campanha - 2006  
Registros de pesquisas eleitorais  
Processos eleitorais entre candidatos  
Prestações de contas - 2006  
Responsáveis com prestações de

Figura 31 - O conteúdo do *blog Deu no Jornal* também pode ser lido no celular. Dez.2006.

Também em alguns outros produtos, pode-se constatar a utilização do material de arquivo integrado no contexto da oferta informativa e não apenas na seção “Arquivo”. A *BBC News* faz isso na seção “On this day”, na qual o material de arquivo é organizado nas subseções ou categorias: “Front page”, “Years”, “Themes”, “Witness” e “Correspondents”. As duas últimas são bem interessantes, pois agregam, de um lado, a participação dos usuários, que podem dar seu depoimento sobre um determinado fato ocorrido do qual tenham sido testemunhas, e, por outro, os relatos dos correspondentes internacionais da *BBC* sobre a cobertura dos eventos. Além de texto, as informações são complementadas por áudios, vídeos, gráficos, entre outros recursos.

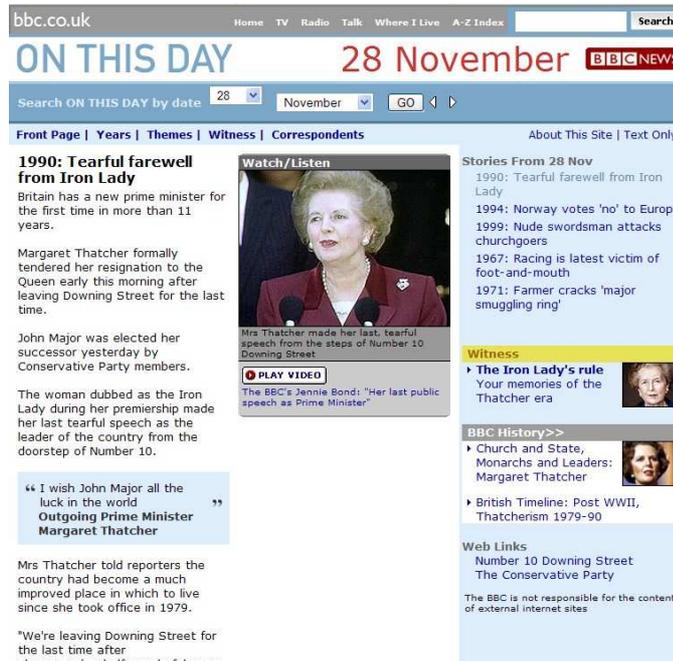


Figura 32 - BBC News agrega as “memórias” de acontecimentos enviados por usuários, além de disponibilizar os relatos dos eventos a partir dos seus correspondentes espalhados pelo mundo. Nov.2006.



Figura 33 – No Libération, o arquivo está integrado às matérias, oferecendo contextualização e aprofundamento. Também se trabalha com o inter-relacionamento. Jan.2007.

Essas são algumas possibilidades que se tem observado para trabalhar o conteúdo jornalístico a partir da utilização das bases de dados. Como se vê, mesmo aqueles produtos de grandes e tradicionais conglomerados de comunicação estão incorporando soluções para o tratamento e processamento das informações, e apresentando-as de maneira mais diferenciada, o que vem alterando a configuração estética dos respectivos *sites*.

No início de janeiro de 2007, o *The New York Times*, por exemplo, disponibilizou um especial sobre as vítimas da guerra no Iraque – “Casualties of War” – agregando informações do Departamento de Defesa dos Estados Unidos sobre cada membro do serviço norte-americano que tem morrido desde o início da ocupação norte-americana, em 2003. Com o subtítulo “Faces of the Dead” – que lembra o do *site* do *Washingtonpost.com*, referenciado mais acima, também sobre os soldados mortos no Iraque –, o conteúdo tem como interface uma “imagem mosaico” composta por pequenos quadrados que formam o rosto de um soldado determinado, focalizado no instante em que se clica para ler sobre o histórico da sua morte.

As informações foram estruturadas e organizadas a partir da data da morte, sendo que as mais recentes estão concentradas nos quadrados superiores à esquerda da imagem. O especial oferece, ainda, a opção para se analisar o *status* demográfico e militar de cada membro, e para se escutar histórias sobre nove jovens soldados e marines que morreram na guerra, a partir da narração feita por outros militares que acompanharam os momentos finais de cada um deles. *Links* relacionados para outros textos, fotos, áudios e material de arquivo complementam o material multimídia.



Figura 34 – *Faces of the Dead* do *The New York Times*: interface “imagem mosaico” para contar histórias dos soldados mortos no Iraque. Jan.2007.

Outras soluções, talvez consideradas como mais simples e que, no entanto, conferem dinamismo e maior interatividade aos produtos, são aquelas de aumentar ou diminuir o tamanho da fonte, escolher visualizar uma matéria em duas ou três colunas, escutar a notícia, ver apenas a versão em texto, ou, o que é mais recente, incorporar opções de compartilhamento de artigos, por meio de aplicações da chamada web 2.0<sup>197</sup>.

<sup>197</sup> Não existe um consenso sobre o uso do termo ou mesmo um conceito claro sobre web 2.0. Há, segundo Tim O’Reilly (2005), que cunhou o termo, um conjunto de princípios e práticas que formam uma espécie de núcleo gravitacional em torno da idéia de web 2.0, proposta como a segunda geração da internet, também chamada “web social”. A disponibilização de aplicações e serviços para compartilhamento do conhecimento e da produção social e livre de conteúdos é um dos eixos da web 2.0. Para o jornalismo, a web 2.0 configura o que alguns autores denominam como Jornalismo 3.0, ou jornalismo participativo (VARELA, 2005). Os serviços e aplicações da web 2.0 requerem, por exemplo, bases de dados muito mais robustas, servidores, entre outros (In: *Wikipedia*, [http://en.wikipedia.org/wiki/Web\\_2](http://en.wikipedia.org/wiki/Web_2)).

Por exemplo, “digg este artigo”, isto é valorá-lo para ser acrescentado no *site* de fonte aberta *Digg*<sup>198</sup>, *Del.icio.us*<sup>199</sup>, criar *link* para o *blog* pessoal, entre outros.

Tais recursos são comandos agrupados no *corpus* das páginas e que permitem modificar a apresentação dos conteúdos, criar vínculos para outras plataformas, dando um pouco mais de “poder” para o usuário interagir com o material jornalístico. A incorporação das opções para compartilhar as matérias e reportagens agregando-os aos chamados “*sites* sociais” é uma estratégia que rende não apenas mais acessos e publicidade, mas, sobretudo, agrega mais valor “social” aos produtos jornalísticos, considerados, em geral, mais fechados à idéia de compartilhamento. Uma visão que, paulatinamente, está sendo modificada.

---

<sup>198</sup> O *Digg* funciona segundo uma tecnologia que permite hierarquizar a ordem da aparição das notícias no seu *site*, organizando-as dentro de categorias como “World”, “Technology”, Business” “Science”, “Videos”, “Entertainment”, e “Gaming”. A posição é definida pelos usuários, que submetem os *links* das notícias, indicando as respectivas fontes originais deles. Ou seja, quanto mais um *link* é indicado maior será também o tráfego gerado para os *sites* jornalísticos ou *blogs*. Daí porque muitos deles já incorporaram a opção “*digg* este artigo”. Os computadores do *Digg* usam fórmulas especiais para analisar quais notícias estão sendo “votadas” ou recomendadas e, a partir disso, a posição delas pode ir se alterando.

<sup>199</sup> <http://del.icio.us/>. Serviço de gestão de marcadores sociais na web. Permite agregar os marcadores que antes se guardava nos favoritos e categorizá-los com um sistema de etiquetas.



Figura 35 – O *The New York Times* também incorporou opções para agregar seu conteúdo a “sites sociais”. Dez.2006.

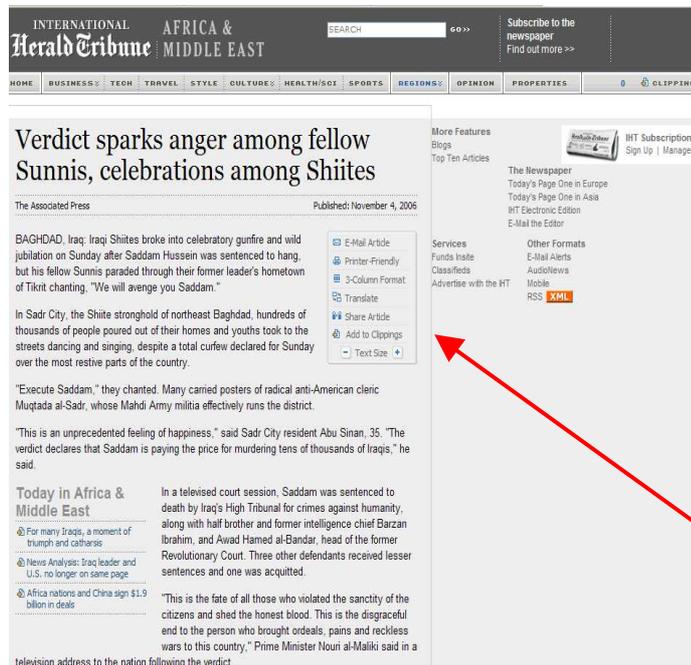


Figura 36 – A seta indica recursos para alterar a configuração do texto, traduzi-lo para idiomas distintos, compartilhá-los com sites da web 2.0, e permitem, também, escolher o número de colunas que se deseja ler o *International Herald Tribune*. As opções dinamizam o produto, que, por outro lado, mantém a referência da metáfora do impresso, com a leitura das matérias em duas ou três colunas. Nov.2006.

Os espanhóis *Elpaís.com* e *Elmundo.es*, além de trabalharem bem com as opções referidas acima, incorporam recursos de inter-relacionamento entre matérias, aplicando efetivamente o *link* para a contextualização dos conteúdos. O *Elmundo.es* é um dos pioneiros no inter-relacionamento para as matérias, remetendo não apenas para o material produzido por suas equipes, mas, também, para produtos concorrentes e em idiomas distintos. Esta opção vem sendo empregada desde fevereiro de 2006, como podemos observar.

The image shows a screenshot of the Elmundo.es website. The main article is titled "CONDOLENCIAS A LA MONJA ASESINADA" and "Benedicto XVI llama al 'respeto mutuo' de las religiones y condena 'toda forma de violencia'". The article text discusses the Pope's statement on religious respect and violence, mentioning the death of a nun in Somalia. To the right of the article, there is a photo of Pope Benedict XVI and a nun. Below the photo, there is a caption: "Fotografía de archivo de la monja italiana Leonela Sgorbati (I) saludando al papa Juan Pablo II en Milán. (Foto: EFE)".

Below the article, there are several sections:

- ADemás:** "Lea el discurso íntegro del Papa" and "Álbumes: Indignación | La polémica, según Al Yazira".
- Divertíos:** A section for entertainment with a search bar and a "Haz clic aquí" button. Below it, it says "¡Encuentra a millones de solter@s en línea!" and "mestic".
- INFORMACIONES ANTERIORES:** "Ahmadineyad expresa su respeto por el Pap".
- Noticias relacionadas en elmundo.es:** "Noticias relacionadas en otros medios".

A red arrow points from the bottom of the article towards the "Noticias relacionadas" section.

Figura 37 – A caixa com as opções de ler mais notícias relacionadas a um determinado assunto seja no *Elmundo.es* ou em outros *sites* é colocada ao final das matérias e no canto direito. Ao clicar, o usuário é levado à página com os resultados indexados. Set. 2006.



Figura 38 – Página mostra lista de *links* para se ler mais sobre, por exemplo, assuntos relacionados ao papa Bento XVI. *Elmundo.es*. Set.2006.

O *Elpais*, seu maior concorrente, oferece os ícones de compartilhamento para os artigos, e, mais recentemente, passou a agregar os *links* para se ler um mesmo assunto em outros *sites* noticiosos, abrindo a possibilidade de se ter o acontecimento tratado em diferentes ângulos, entre outras opções, como a de comunicar erros, ver estatísticas de acesso para uma determinada peça informativa, entre outras.



Figura 39 - *Elpais.com*. Seta indica caixa com opções de inter-relacionamento. Abaixo, os ícones para compartilhamento. Dez.2006.

São soluções que, certamente, despertam maior interesse, além do que, aprimoram as opções de interatividade do produto, demandando mais ações para serem executadas pelos usuários. Eles devem clicar sobre determinados ícones para ativar tais recursos e, em outros casos, para personalizar os conteúdos segundo as suas necessidades. Com isso, os produtos tornam-se uma experiência mais envolvente e certamente mais autêntica. “Significa uma nova repartição do trabalho cognitivo entre o *site* e seus usuários”, assinalam Francesco Cara e Stephana Broadbent (2003, p.206).

Embora ainda exista defasagem entre aquilo que a tecnologia oferece e a incorporação dela por parte de empresas jornalísticas – por questões culturais e de investimento, entre outras – percebe-se que a adoção de BDs, além de recursos visando a uma maior dinamização e qualificação dos *sites*, dos respectivos conteúdos oferecidos, além das soluções aplicadas para automatizar os processos de apuração, tratamento editorial e apresentação de conteúdos – de modo parcial ou total –, é uma decisão estratégica para a implementação dos produtos. Mais cedo ou mais tarde, empresas de perfis distintos certamente irão aderir, em maior ou menor grau. Pois, como escreveu Nilson Lage (2002), quem decide desenvolver e incorporar tecnologia para melhorar o funcionamento dos *sites* jornalísticos estará à frente dos concorrentes e no ritmo de seu tempo. Certamente, estas são ou serão “empresas e instituições interessadas em tomar a chave, rodar o mecanismo e abrir a porta para o futuro”.

E, finalizando o tópico, lembramos com Mindy McAdams (2005, p.440) que, desde 2004, usar bases de dados para gerenciar *web sites* jornalísticos tem se tornado uma norma. Stephen Quinn (2002, p.128) as compreende como uma das implementações fundamentais para gerir o conhecimento nas redações digitais. Uma vez empregadas com sabedoria, as bases de dados aliadas a linguagens como a *XML*, podem ser utilizadas para definir *tags* (marcas, etiquetas) para facilitar o acesso às

informações estruturadas e distribuí-las em formatos diversos, além do que, como sustenta Quinn (2005, p.143), se tornam estratégia efetiva para a contextualização e complementação das informações jornalísticas.

### 3.1.3 Sistemas de gestão de conteúdos e as bases de dados

Os sistemas de gestão de conteúdos são as chamadas ferramentas ou sistemas de publicação, que visam, principalmente, facilitar o processo de redação, composição, edição e apresentação de todas as peças informativas, imagens fixas e em movimento, áudios, infográficos, bem como a sua disponibilização pelos produtos digitais<sup>200</sup>. Tais sistemas, desenvolvidos especialmente para as empresas que mantêm produtos jornalísticos digitais ou por suas próprias equipes de tecnologia, são elaborados por profissionais de informática e se constituem em soluções, isto é, *softwares* para otimizar o trabalho do jornalista e reduzir a necessidade de que ele tenha conhecimentos técnicos sobre o processo de produção e disponibilização do material na web (MIELNICZUK, 2003).

A criação de sistemas de gestão de conteúdos mais compatíveis com as características do jornalismo digital e mais afinados com as necessidades jornalísticas passou a ser mais efetiva há pouco mais de três anos, de modo mais sistemático. Antes

---

<sup>200</sup> Embora estejamos nos referindo especificamente às ferramentas de publicação no jornalismo, a grande maioria dos *sites* utiliza tais ferramentas, mais convencionalmente denominadas *Content Management Systems (CMS)*, para publicar conteúdos de qualquer natureza na web. Os *blogs*, os conhecidos diários da rede, são também um exemplo de ferramenta de publicação. Neste caso, ela é constituída de um formulário com campo para inserção do texto, dos *links*, das imagens, e com botão para publicar o referido material. De certa forma, os *blogs* expandiram a utilização desses sistemas, automatizando e facilitando a publicação de conteúdos na web.

disso, em geral, as empresas utilizavam ferramentas-padrão desenvolvidas comercialmente e que nem sempre atendiam aos propósitos do jornalismo<sup>201</sup>.

Um dos fatores que têm contribuído para a melhor adequação de tais sistemas é a incorporação de jornalistas trabalhando de modo articulado com os programadores, desenvolvedores, entre outros profissionais de informática (MIELNICZUK, 2003; MIELNICZUK; MARQUES, 2006; LÓPEZ, GAGO, PEREIRA, 2003; LÓPEZ, 2006; GAGO; 2006; SCHWINGEL, 2003, 2004, 2005; SCHWINGEL; MACHADO; PALACIOS; ROCHA, 2005; MACHADO, 2006).

Carla Schwingel (2004, p.05) esclarece que um sistema de publicação visa “incorporar efetivamente as características do jornalismo digital, tanto na concepção do sítio (na arquitetura da informação do produto), quanto na estrutura da notícia (na arquitetura da informação de cada matéria)”. Ela faz, inclusive, uma diferenciação entre ferramentas de publicação de conteúdo como o *Content Management System (CMS)*, que é utilizado para selecionar, aprovar, editar e publicar, e o *Portal System*, utilizado para facilitar a manipulação dos módulos do *site*.

No jornalismo digital, os dois tipos são empregados, porém, com a incorporação de recursos para integrar a hipertextualidade, a multimídia, a interatividade, a personalização, o material de arquivo e a característica da atualização contínua. Sendo assim, a autora define os sistemas de publicação como: “Ambientes tecnológicos que possuem as potencialidades dos sistemas de administração de conteúdos e de gerenciamento de portais, adaptados às rotinas produtivas do jornalismo digital (SCHWINGEL, 2004; 2005). Ou seja, integrando:

---

<sup>201</sup> Pudemos constatar isso quando realizamos a pesquisa de mestrado sobre o jornalismo digital em portais regionais (com estudo de caso sobre o *UAI*, [www.uai.com.br](http://www.uai.com.br), e o *iBahia.com*). Naquele momento, uma das limitações e mesmo entraves relacionados à produção de conteúdos, segundo reportado pelos profissionais daqueles produtos, era justamente a inadequação das ferramentas de publicação disponíveis (BARBOSA, 2002; 2003).

- Ferramentas que explicitem o processo de apuração de uma matéria;
- Ferramentas que facilitem a publicação das informações de forma a apresentar modelos previamente elaborados de roteiros narrativos para os distintos gêneros jornalísticos (os informativos, como a notícia; os interpretativos, como a reportagem; os dialógicos, como a entrevista, o fórum, o *chat*, a enquete; argumentativos, como os comentários, as colunas, os artigos, os editoriais; e a infografia<sup>202</sup>); e
- Ferramentas que facilitem a veiculação e distribuição das informações (SCHWINGEL, 2005, p.10).

Como um dos pilares balizadores dos sistemas de gestão de conteúdos estão as bases de dados: “Una base de datos ya no es solo una aplicación utilitaria: es un vehículo de canalización de la producción periodística, es un condicionante de como se representa la información en pantalla, es la clave para generar hipertexto efectivo y rápido” (LÓPEZ, GAGO, PEREIRA, 2003, p.195)<sup>203</sup>.

E, associadas ao emprego das bases de dados para a criação dos sistemas de gestão de conteúdos, estão também as linguagens de programação (MIELNICZUK; MARQUES, 2006, p.03): “As linguagens de programação dedicadas ao desenvolvimento de sítios dinâmicos (entre as mais utilizadas estão o PHP, o ASP, o

---

<sup>202</sup> Há uma discussão estabelecida em torno dos gêneros jornalísticos e como eles são reconfigurados, resultando em híbridos ou mesmo novos gêneros com padrões e formatos editoriais bastante consolidados no jornalismo digital: a crônica de última hora ou *flash* (*hard news* e *breaking news*), a infografia multimídia e interativa, os *blogs* são alguns citados por Ramón Salaverría (2005b, p.144), por exemplo. Não é nosso objetivo aqui discutir o assunto, apesar de concordamos com os autores (DÍAZ NOCI, 2001; DÍAZ NOCI; SALAVERRÍA, 2003; SANCHO, 2003; MARCOS, 2003; SALAVERRÍA, 2005a; SALAVERRÍA, 2005b; EDO, 2002, 2005; RIBAS, 2005; SEIXAS, 2004; 2006) que consideram o potencial dos *chats*, fóruns, da infografia, dos *blogs* - e, segundo a nossa abordagem, também os arquivos - como possíveis novos gêneros.

<sup>203</sup> “Uma base de dados já não é só uma aplicação utilitária: é um veículo de canalização da produção jornalística, é um condicionante de como se representa a informação na tela, é a chave para gerar hipertexto efetivo e rápido”. (t.A.).

Java e o Coldfusion) oferecem códigos que podem ser mesclados aos códigos HTML, possibilitando inserir comandos ao sítio”. Outra técnica que vem sendo bastante utilizada é a *Ajax (Asynchronous Javascript and XML)*<sup>204</sup>. Para desenvolver a ferramenta de publicação para o novo *Urbi et Orbi*, por exemplo, Marco Oliveira, responsável técnico do *site*, empregou a *Ajax*, associada à PHP (*Hypertext Preprocessor*)<sup>205</sup> e a bases de dados relacionais *Postgree*, entre outros *softwares* livres<sup>206</sup>.

Ao ministrar a Oficina de Jornalismo Digital, do 6º semestre do curso de Jornalismo da FACOM/UFBA, Carla Schwingel (2003), ao lado de Eduardo Miranda (2006), aluno da oficina e programador, empregou a linguagem *Active Server Page (ASP)*<sup>207</sup>, e também bases de dados relacionais, para criar o *Produto de Implementação do Panopticon (PIP)* para o produto laboratorial homônimo da disciplina. O *Panopticon* foi idealizado pelo professor Marcos Palacios em 2001<sup>208</sup>, e, com o *PIP*, desenvolvido durante o segundo semestre de 2002, o produto passou de um *site* estático, que usava o *Dreamweaver* para a sua atualização, para se tornar um *site* dinâmico<sup>209</sup>.

---

<sup>204</sup> A *Ajax* não é exatamente uma tecnologia em si, mas, um termo que se refere ao uso de um grupo de tecnologias para criar aplicações web interativas. Ela proporciona às páginas web trocarem pequenas quantidades de dados com o servidor e, assim, essas páginas não têm que ser recarregadas cada vez que o usuário fizer alguma mudança (*Wikipedia.com*). Por exemplo, o *Google* usa a *Ajax* no *Gmail*, o que permite que os novos emails que cheguem à caixa-postal sejam mostrados instantaneamente, sem precisar que se clique em qualquer botão para saber se uma nova mensagem chegou.

<sup>205</sup> Ambiente de programação desenvolvido pela comunidade de *software* livre.

<sup>206</sup> Informação coletada em entrevista concedida à autora no dia 05/06/2006.

<sup>207</sup> *ASP* é um ambiente de programação desenvolvido pela *Microsoft* para elaboração de páginas web dinâmicas. Ela mescla HTML com outras linguagens de programação. Seu principal concorrente é a *PHP*.

<sup>208</sup> Tivemos a oportunidade de acompanhar a primeira fase do *Panopticon*, atuando como aluna de tirocínio docente à época do Mestrado, junto ao professor Marcos Palacios e Gottfried Stockinger (que, com o afastamento do professor Palacios para o pós-doutorado, ficou como responsável pela disciplina).

<sup>209</sup> Entre o *Dreamweaver* (programa da *Macromedia*) e o *PIP*, também se trabalhou com a ferramenta *blog*.

O PIP foi proposto com o intuito de publicar de forma automatizada e facilitada (...) Constitui-se em um sistema dinâmico que visa à divulgação de arquivos em banco de dados. Sua elaboração ocorreu em ambiente ASP integrado com o banco de dados SQL Server (...) Projeto e sistema foram discutidos a cada aula, de acordo com o avanço da teoria. E a partir da compreensão das características do Jornalismo Digital, suas aplicabilidades no *Panopticon* foram sendo definidas (SCHWINGEL, 2003, p.194).

Em 2004, o *PIP* transformou-se na *PPEP – Plataforma de Publicação e Ensino Panopticon* –, um protótipo que vem sendo desenvolvido em bases de dados e que aplica teorias jornalísticas dentro de um novo ambiente de publicação e ensino na internet (SCHWINGEL; MACHADO; PALACIOS; ROCHA, 2005). A *PPEP* está sendo criada como um subproduto da Rede Latino-Americana para o Ensino do Jornalismo, envolvendo o Grupo de Pesquisa em Jornalismo On-line da FACOM/UFBA, a Universidade Federal de Pernambuco (UFPE), a Universidade de Santa Maria (UFSM/RS), a Universidade Federal de Santa Catarina (UFSC), a Universidade de Córdoba (Argentina), sendo que em junho deste ano as Faculdades Jorge Amado (FJA/BA) também foram integradas ao projeto. A plataforma está estruturada como um ambiente multiusuário para ser utilizada simultaneamente em uma ou mais instituições e faz parte da pesquisa doutoral de Carla Schwingel<sup>210</sup>. O protótipo está sendo desenvolvido com *software* livre e de fonte aberta, empregando *PHP*, banco de dados *Postgree* e o pacote de *software PEAR*.

Em dezembro de 2006, a equipe de desenvolvedores estava trabalhando na hospedagem da solução nos servidores das Faculdades Jorge Amado, e realizando a documentação do produto a fim de efetivar as funcionalidades associadas aos *templates*

---

<sup>210</sup> Em 2005, a doutoranda fez a apresentação dos resultados preliminares durante a V Bienal Iberoamericana de la Comunicación, em Atizapán de Zaragoza, no México.

(modelos) definidos para os produtos jornalísticos que serão gerados<sup>211</sup>. Eles incluem formatos de diferentes tipos, desde jornais, a portais, revistas, *blogs*, entre outros que poderão ser agregados. Vale ressaltar que os *templates* incorporam as características do jornalismo digital e os distintos gêneros jornalísticos.

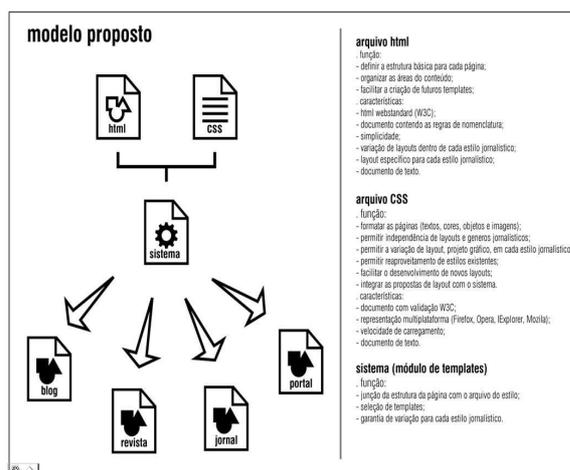


Figura 40 – Modelo proposto para criar produtos com formatos distintos, a partir de *templates*, no novo *Panopticon*. (Dez.2006).

Para a criação desses modelos também estão sendo utilizadas as Folhas de Estilo em Cascata ou *Cascading Style Sheets (CSS)*<sup>212</sup>, documento onde são definidas regras de formatação ou de estilos a ser aplicadas aos elementos estruturais de um sistema. Com a *PPEP*, as etapas do processo de produção jornalístico estarão integradas:

<sup>211</sup> As informações sobre essa fase mais atual nos foram fornecidas por Beatriz Ribas, integrante do GJOL e membro da equipe que responde pela plataforma, por parte das Faculdades Jorge Amado, onde é professora da disciplina Jornalismo Digital.

<sup>212</sup> As *CSS* permitem que seja retirada da marcação HTML toda a formatação (apresentação) do documento web. Quem vai determinar cores, formas, tipos e tamanhos, posicionamentos em todo o visual da página são as *CSS*.

– Apuração. Já prevê a construção de pautas hipertextuais, organização e contatos com fontes nominais (contactadas via *e-mail* ou telefone) e no ciberespaço (com a indicação precisa do *site* ou *sites* que contêm a informação);

– Composição das peças informativas. Envolvendo o processo descentralizado de elaboração da narrativa hipertextual, com níveis de aprofundamento para o usuário que permitam estruturação dos blocos de texto em camadas, numa construção fragmentada, e a visualização de um discurso único e coerente;

– Circulação. Envolvendo novas operações de uso relacionadas à arquitetura da informação<sup>213</sup> e à usabilidade.

Outro sistema publicador, ou melhor protótipo, criado no âmbito acadêmico e que aplica a base conceitual do jornalismo digital, é o *MapaLink*<sup>214</sup>, um projeto desenvolvido por Luciana Mielniczuk em conjunto com o então aluno de Jornalismo na UFSM, e programador Iuri Lammel Marques, em 2005. O *MapaLink* é o protótipo de um sistema publicador que permite desenvolver narrativas jornalísticas hipertextuais, sem exigir do jornalista conhecimentos técnicos de programação (MIELNICZUK; MARQUES, 2006, p.10). Assim, ele permite:

- Segmentar a narrativa em blocos de texto e organizá-la em uma estrutura de microsítio, isto é, elaborar cada matéria como um microportal, segundo apresentado por Carla Schwingel (2004);

- Gerar automaticamente um menu da notícia;

---

<sup>213</sup> Concebida como um roteiro audiovisual que baliza a produção (SCHWINGEL, p.08). Consiste na explicação tela a tela dos possíveis desdobramentos da matéria, tendo o cuidado com a linha narrativa. Por exemplo, o texto para a composição da pauta, será composto na própria estrutura do sistema publicador (SCHWINGEL; MACHADO; PALACIOS; ROCHA, 2005, p.110).

<sup>214</sup> <http://coralx.ufsm.br/facos/mapalink//admin.php>.

- Criar automaticamente uma representação gráfica da narrativa hipertextual, a partir de um mapa interativo.

A ferramenta oferece ao jornalista alguns diagramas que representam os *links* para se estruturar as matérias, que são apresentadas nas seções “Política”, “Economia”, “Entretenimento” e “Esportes”. São eles: “Acontecimento”, “Complemento” (representa os *links* de detalhamento, de exemplificação ou particularização e de complementação ou ilustração), e “Versões” (representa o *link* de oposição). Desta forma, como explicam Luciana Mielniczuk e Iuri Marques, o jornalista pode clicar em um dos diagramas oferecidos para que a estrutura da narrativa seja modificada, de acordo com o tipo de informação do segmento que continuará a seqüência (MIELNICZUK; MARQUES, 2006, p.12).

**Congresso apressa votação da reforma**  
22/11/05

O Congresso Nacional está disposto a adiar a votação do projeto de reforma, proposto por deputados da bancada governista. Segundo o presidente da Câmara dos Deputados José da Silva, grande parte dos deputados já anunciou a pretensão de apressar a reforma...

**Continuação da narrativa**

- ▶ [Oposição votará contra o projeto](#)
- ▶ [Governo diz que jogará pesado para garantir aprovação.](#)

**Leia também**

- [Prefeituras municipais enfrentam problemas financeiros](#)
- [Presidente visita países do Oriente Médio](#)
- [Vereadores fecham Câmara em cidade do interior](#)

**Congresso apressa votação da reforma**

- ▶ **Oposição votará contra o projeto**
  - ▶ João Silva, presidente do PCP, critica projeto de reforma
  - ▶ Maria Machado, do PCP, contesta adiantamento da votação.
- ▶ **Governo diz que jogará pesado para garantir aprovação.**
  - ▶ Senador pereira, líder da bancada governista, comemora
- ▶ **O que muda com a aprovação da reforma**

**Mapa da narrativa**

Figura 41 – Página do publicador *MapaLink* com a representação da notícia hipertextual.  
Fonte: MIELNICZUK; MARQUES, 2006, p.14.

Para o usuário, o *MapaLink* proporciona parâmetros para que ele possa guiar-se pela navegação do microsítio, enquanto elege seu roteiro de leitura de cada bloco/camada que compõe as peças informativas. O protótipo também será agregado à plataforma *PPEP*, conforme previsto no planejamento. Uma observação pertinente a ser feita sobre este protótipo é quanto à inserção de imagens, de áudio, entre outros recursos multimídia. Embora seja dito que ele não tem o objetivo imediato de disponibilizar tais

recursos, seria recomendável incorporá-los no seu diagrama de complemento ou em outro específico. Do contrário, pode ficar limitado, impedindo que também se ampliem as possibilidades para a construção de peças informativas mais diferenciadas, e afinadas com as propriedades do jornalismo digital.

À semelhança de qualquer sistema de gestão de conteúdos, o *MapaLink* possui níveis diferenciados de permissão para os usuários. Ou seja, o administrador, em geral, tem uma senha que lhe autoriza o acesso a todas as áreas do sistema, os editores, têm acesso a toda a parte relativa ao conteúdo, enquanto repórteres têm um nível de permissão que lhes dá acesso às telas do sistema para a composição, e edição, não sendo liberado o acesso para modificar, por exemplo, a *home* de um produto.

Claro que os níveis de permissão podem variar de acordo com o sistema e com o produto, com os profissionais que fazem parte das equipes, além dos colaboradores e dos próprios usuários. Por exemplo, no sistema do *Urbi et Orbi*, qualquer usuário pode inserir uma matéria, conforme habilitado pela ferramenta. Porém, antes da publicação efetiva dela, o material é filtrado por um dos diretores do jornal, através de um comando que foi inserido na ferramenta. De modo geral, os diversos *sites* jornalísticos que incorporam opções para a participação dos usuários, seja enviando textos, fotos, vídeos, comentários, ou qualquer outro material, estabelecem algumas normas para adequá-los às respectivas linhas editoriais. O envio também pode estar condicionado ao respeito a determinadas leis dos países (*Elpais.com*), à uma idade mínima de 13 anos, por exemplo, no caso do *Topix.net*, ou de 18 anos e cessão dos direitos à empresa, como estipulado pelo *Portal Terra*.

Já a *BBC News* definiu duas formas de moderação para a seção “Have Your Say”<sup>215</sup>. A depender do assunto pautado para a discussão, a participação pode ser totalmente moderada, isto é, *Fully Moderated*, ou reativamente moderada, *Reactively Moderated*. Nesta segunda modalidade, caso o usuário seja registrado no *site*, seu comentário será publicado diretamente sem filtragem. Até mesmo os produtos de fonte aberta (*open source*), como o *Independent Media Center*<sup>216</sup> e suas versões, incluindo a brasileira *Centro de Mídia Independente*<sup>217</sup>, possuem algumas regras para que o material dos usuários seja publicado nos seus respectivos *web sites*<sup>218</sup>.

### 3.1.4 Tipos de bases de dados

Destacamos neste subtópico algumas especificações técnicas sobre as bases de dados, pois, a partir delas, compreende-se melhor as distintas classes, modelos ou tipos de BDs. Antes de mais, vale notar com Dan Brincklin (2001) que as bases de dados, freqüentemente, podem ser mantidas e organizadas manualmente ou num computador,

---

<sup>215</sup> <http://newsforums.bbc.co.uk/nol/thread.jspta?threadID=5096&start=0&&edition=2&ttl=20070102193229>. Esta é mais uma seção que a *BBB News* abriu à participação do usuário, pois em “On this day” também é oferecido espaço para o público registrar seus testemunhos sobre eventos passados.

<sup>216</sup> <http://www.indymedia.org/en/>.

<sup>217</sup> <http://www.midiaindependente.org/>.

<sup>218</sup> Em outros *sites* de fonte aberta, como o *Digg*, é recomendado ao usuário observar algumas regras para manter o bom funcionamento e a qualidade do conteúdo publicado. No *Slashdot* ([www.slashdot.com](http://www.slashdot.com)), é solicitado aos usuários que se cadastrem, e seus artigos passam por um processo de seleção para avaliar o nível de interesse, atualidade e relevância. No *OhmyNews* (<http://english.ohmynews.com/>), um dos exemplos mais citados quando se fala de jornalismo participativo ou de jornalismo cidadão, os registros dos usuários interessados em se tornarem “repórteres cidadãos” são avaliados por um editor e, após liberada a participação, devem seguir o “código de ética”, composto por oito itens, sendo o primeiro deles: “The citizen reporter must work in the spirit that “all citizens are reporters,” and plainly identify himself as a citizen reporter while covering stories”. (“O cidadão repórter deve trabalhar no espírito de que “todos cidadãos são repórteres”, e identificar claramente ele próprio como um cidadão repórter enquanto cobre os eventos” (t.A.).

de modo automatizado. Brincklin (2001) se refere ainda às BDs manuais voluntárias, nas quais os usuários introduzem os dados a partir de seu interesse pelo tema, compartilhando com outros, ou, então, criam suas próprias BDs. Este, por exemplo, é o modelo utilizado por boa parte das BDs multimídias (QUIJADA, 2006), como a *Compact Disc Data Base (CDDB)*<sup>219</sup>.

Em geral, quando as BDs possuem muitos tipos diferentes de dados (textos, imagens fixas ou em movimento, infográficos, áudios, etc.) interdependentes e inter-relacionados, como no caso de um produto jornalístico digital, elas são consideradas como complexas, requerendo um trabalho minucioso para estruturar e indexar adequadamente registros de tipos tão diversos. Pelo que notamos, as bases de dados relacionais são as mais empregadas para a organização e a estruturação das informações no jornalismo e no jornalismo digital, ainda que possam estar associadas a outros modelos.

Como vimos no capítulo anterior, as BDs relacionais são aquelas que possuem duas ou uma coleção de muitas tabelas, planejadas para ser usadas em associação uma com a outra. As BDs relacionais permitem comparações entre duas ou mais tabelas e, portanto, também os cruzamentos entre os registros inseridos, ou seja, as informações contidas neles. Mas, além delas existem outros modelos de BDs, segundo a literatura das ciências da computação. A seguir, descrevemos esses modelos a partir da tipificação elaborada pelos autores Célio Guimarães (2003) e Lee Jorgensen (2003):

---

<sup>219</sup> [www.gracenote.com](http://www.gracenote.com). Cada vez que se copia um CD, a *CDDB* informa *online* os dados referentes àquele disco (nome, intérprete, a lista de canções, entre outros). Esta aplicação é encontrada em reprodutores como *Winamp*, *Musicmatch* ou *Real Jukebox*.

- Hierárquico ou em árvores: A informação é organizada em níveis;
- Em rede: Nele, cada registro mestre é acessado por uma chave e, a partir dele, vai-se seqüencial ou diretamente aos registros-membro;
- Relacional: Tipo mais popular. Os dados e as relações entre eles estão organizados em tabelas formalmente chamadas de relações e, alguns campos, podem ser designados como chaves, o que significa dizer que a procura por certo tipo de informação pode ser acelerada utilizando as indexações criadas com as chaves;
- Orientado a Objetos: Mais avançado do que as relacionais. É atualmente um dos principais tipos, pois são mais adequados para o tratamento de objetos complexos (textos, gráficos, imagens) e dinâmicos (programas, simulações). Combina idéias dos modelos de dados tradicionais e de linguagens de programação orientadas a objetos. Noção de objeto é usada no nível lógico e possui características não encontradas nas linguagens de programação tradicionais, como operadores de manipulação de estruturas, gerenciamento de armazenamento, tratamento de integridade e persistência dos dados. As BDs relacionais têm incorporado recursos orientados a objetos;
- Sistemas de BDs Inteligentes (*Intelligent Database Systems*, IDBS): Aqueles derivados da integração da tecnologia de BDs com técnicas desenvolvidas no campo da inteligência artificial (IA). Eles representam a evolução e a união de diversas tecnologias, entre elas a hipermídia<sup>220</sup>.

---

<sup>220</sup> A hipermídia é definida por Lúcia Leão (1999) como uma tecnologia que engloba recursos do hipertexto e multimídia, permitindo ao usuário a navegação por diversas partes de um aplicativo, na ordem que desejar (LEÃO, 1999, p.16). Já Vicente Gosciola (2003), que vê a hipermídia como um meio, uma linguagem e um produto audiovisual, sendo a sua matriz o audiovisual e não o hipertexto, a define como um conjunto de meios que permite acesso simultâneo a textos, imagens e sons de modo interativo e não-linear, possibilitando fazer *links* entre elementos de mídia, controlar a própria navegação e, até, extrair textos, imagens e sons cuja seqüência constituirá uma versão pessoal desenvolvida pelo usuário (GOSCIOLA, 2003, p.34).

Encontramos, ainda, outras especificações sobre classes de bases de dados (ativas, dedutivas e multimídias) e que podem ser incorporadas às BDs relacionais, resultando, assim, numa boa combinação para o emprego no jornalismo digital. Vejamos, então, quais são as definições para elas, de acordo com Mário Piattini e Coral Calero (1999), Mário Piattini (2000), Sílvia Salgado e Rosendo Rodriguez (2003), Paulo Villegas e Daniel Garcia (2002) e David Fernández Quijada (2006):

- BDs Ativas. Ao contrário das tradicionais BDs passivas, as BDs ativas se produzem em certas condições em que há a necessidade de se executar determinadas ações de modo automático. Elas devem ser capazes de monitorar e de reagir a eventos de maneira oportuna e eficiente. Garantem maior produtividade, possibilidade de otimização semântica e facilita o acesso dos usuários finais às BDs;

- BDs Dedutivas. Têm a capacidade de definir regras com as quais vai deduzir ou inferir informação adicional, a partir dos dados armazenados nas BDs. Uma base de dados dedutiva é um programa lógico, para mapear relações base a partir dos fatos, e regras que são usadas para definir novas relações em termos das relações base e termos de consultas;

- BDs Multimídias. As BDs multimedia têm experimentado rápido desenvolvimento nos últimos anos, em função da internet e de outras ferramentas telemáticas. As BDs multimídia são classificadas como: referenciais, com as informações sobre o tipo de produto ou objeto que armazena, além dos respectivos títulos, autores, etc. São empregadas para gerir conteúdos tanto em entornos fechados (uma emissora de televisão, por exemplo) como em redes abertas como a internet; e descritivas, vão mais além daqueles dados técnicos, tratando da análise do conteúdo em

si, com descrição dos elementos contidos nos objetos multimídia, por exemplo. Ainda se encontram em estágio embrionário, conforme David Fernández Quijada<sup>221</sup>.

A persistência ou durabilidade dos dados em um meio de armazenamento confiável é um dos requisitos mais básicos de uma BD. “O sistema deve ser suficientemente robusto para se recuperar de uma falta de energia ou mesmo de uma falha nos discos de armazenamento sem perda dos dados” (GUIMARÃES, 2003, p.21). A esse conjunto de requisitos, denomina-se ACID: de atomicidade, consistência, isolamento e durabilidade. O requisito de atomicidade, por exemplo, irá estabelecer a dependência de que certas operações sobre os dados devem ser feitas de forma conjunta e indivisível para preservar a consistência do sistema, mesmo havendo falhas no equipamento ou na comunicação com a base de dados.

Por conta dos requisitos exigidos e da complexidade dos dados ou objetos inseridos numa base de dados, é necessária uma coleção de programas que permita a sua criação e manutenção. Essa coleção de programas é chamada de Sistema Gerenciador de Base de Dados (SGBD) e, entre as suas tarefas, estão: descrever de alguma forma a estrutura dos dados de cada BD específica; permitir apenas acesso autorizado aos dados, sendo que alguns deles só estarão acessíveis a determinada categoria de usuário; prover recursos para o *backup* da base de dados.

Diversas tecnologias de SGBD e linguagens foram desenvolvidas por empresas de *software* e de *hardware* - grandes e pequenas - ao longo dos últimos 30 anos, e, entre elas, uma das mais empregadas e mais bem sucedidas comercialmente é a *SQL* (*Structured Query Language*) ou Linguagem de Consulta Estruturada, da IBM, que se

---

<sup>221</sup> Ele defende a criação de um padrão para as BDs multimídia, nos moldes do sistema ISBN usado no setor editorial, que desde a década de 70 se tornou um modelo de identificação único. Quijada propõe algo semelhante, que seja possível de ser “assinável por via eletrônica e com valor global”.

tornou o padrão dos SGBDs relacionais. Além disso, tal linguagem inspirou o desenvolvimento de outros produtos, a exemplo do *software* livre *MySQL*, atualmente um dos SGBDs mais populares. Ele possui boa integração com *PHP*, sendo utilizado em diversas aplicações, e em *web sites* de referência, como a *Wikipedia*, entre outros.

A especificação técnica de alguns aspectos das bases de dados, como realizado neste subtópico, tem por objetivo dar suporte à proposta de definição e classificação das BDs no jornalismo digital, discutidas a seguir.

O aporte teórico trabalhado neste capítulo sobre o paradigma que as bases de dados instauram no jornalismo digital, assim como os aspectos técnicos das BDs, de acordo com as formulações provenientes do reino das ciências da computação, nos permitiram encontrar a fundamentação necessária para delimitar as particularidades do **Jornalismo Digital em Base de Dados (JDBD)**. Assim, a seguir, apresentamos o conceito que colocamos em operação, identificando as funcionalidades, bem como as categorias que conferem especificidade a este modelo.



## **DELIMITANDO O “JDBD” E AS SUAS PROPRIEDADES**

## 4

**Delimitando o “JDBD” e as suas propriedades**

---

*La tendencia actual privilegia una información cada vez más estructurada, más pautada y organizada, que es posible gracias a la implantación de sitios web `dinámicos`.*  
- Ramón Salaverría; Rafael Cores.

*Las tendencias apuntan pues a un camino de experimentación, innovación y redefinición de los productos que conocemos hasta ahora.*  
- Xosé López.

**4.1 DIMENSÃO CONCEITUAL E OS ELEMENTOS DE ESPECIFICIDADE**

Para entender o que se considera como nova noção de bases de dados aplicada ao jornalismo digital, segundo a nossa abordagem, já não bastam os conceitos de depósito integrado de dados, coleção de documentos ou de repositório de informações para consulta e recuperação. Buscando ampliar a compreensão sobre essa tecnologia da informação e seu emprego no jornalismo digital nessa fase de transição entre a terceira e a quarta geração, recorreremos ao conceito de base de dados formulado por Lev Manovich (2001), que as vê como uma forma cultural estruturadora dos produtos da nova mídia (ou dos produtos digitais) na contemporaneidade, e aquele pensado por Elias Machado (2004a; 2004b; 2006), de base de dados como um formato no jornalismo digital.

Ao incorporar tais concepções acreditamos estar trabalhando com uma idéia de ruptura (PALACIOS, 1999, 2002a, 2003b, 2003c, 2004; MIELNICZUK, 2003; SILVA JÚNIOR, 2006). Tanto em relação ao significado de base de dados, ao uso que se fazia das BDs no jornalismo até então como ferramentas para a melhoria da qualidade das reportagens, como em relação às novas funcionalidades percebidas para a estruturação, a organização e a apresentação das informações jornalísticas, para a construção das peças informativas, para o armazenamento dos conteúdos publicados e a sua inclusão na oferta informativa, enfim, para a gestão dos produtos digitais.

Paralelamente, este cenário nos conduz a pensar as bases de dados como agentes com capacidade de produzir uma nova metáfora (BARBOSA, 2004a, 2004b, 2004c, 2005a, 2005b, 2005c, 2006a, 2006b), superando aquela do impresso, também conhecida como *broadsheet metaphor* (MCADAMS, 1995; WATTERS et al, 1997; 2000; AKESSON, 2003; ERIKSEN & IHLSTRÖM, 2000; IHLSTRÖM; LUNDBERG, 2004) que continua sendo a grande referência e a mais utilizada pelos diferentes tipos de *sites* noticiosos.

Assim, em conformidade com tal panorama conceitual, julgamos pertinente considerar as bases de dados com potencial de fato para gerar um jornalismo digital mais diferenciado, inteligente e dinâmico. Sendo assim, propomos como paradigma o que conceituamos como **Jornalismo Digital em Base de Dados (JDBD)**. A incorporação da tecnologia de BDs segundo significação mais inovadora para a gestão de produtos jornalísticos digitais assegura ambientes – redações - inteligentes e também mais complexos.

Um outro conceito que se encaixa na concepção mais ampla de bases de dados como uma forma cultural para produtos digitais (MANOVICH, 2001), como formato (MACHADO, 2004a, 200b, 2006), paradigma e metáfora (BARBOSA, 2004a, 2004b,

2004c, 2005a, 2005b, 2005c, 2006a, 2006b) para o jornalismo digital é o de resolução semântica (FIDALGO, 2003, 2004). Tal concepção abarca as características de interatividade, hipertextualidade, atualização contínua, memória, multimídia e personalização, assim como nos auxilia a explicar as mudanças para a estruturação, a organização e a apresentação das informações. Resolução semântica também é particularmente importante para a definição dos elementos de especificidade propostos para o modelo **JDBD**, que são delineados neste capítulo.

Além disso, ele abrange uma outra noção intrínseca às bases de dados: a noção de metadados - os dados que fazem referência a outros dados. Ferramentas que são fundamentais para a categorização dos conteúdos, bem como para a estruturação das informações, e que guiam os usuários aos dados, tanto para encontrar informação pontual, como para extrair informação sobre o conjunto, provendo um contexto de grande importância para uma melhor interpretação e contextualização das informações jornalísticas.

Desta maneira, as co-relações entre as notícias inseridas numa base de dados, considerando-se a classificação por meio de diversos campos, bem como as possibilidades combinatórias entre elas permitirão produzir, extrair novas informações, novos dados, que vão gerar mais conhecimento, mais contexto, sobre os eventos. Conectada à noção de metadados está a de metainformação (COLLE, 2002, 2005a, 2005b), e que, juntas, nos permitem classificar os produtos digitais estruturados em bases de dados como metainformativos.

O gerenciamento do conhecimento nas redações, de modo geral, e das redações de produtos digitais, de modo particular, depende da incorporação de bases de dados (QUINN, 2002, 2005; COLLE, 2002, 2005a, 2005b; LÓPEZ, GAGO, PEREIRA, 2003; QUADROS, 2004, 2005; MACHADO, 2004a, 2004b, 2004c, 2006; MCADAMS, 2005;

HOLOVATY, 2005, 2006a, 2006b; GAGO, 2006), assim como a oferta de conteúdos dinâmicos, mais contextualizados, também passa pela adoção de BDs.

Cabe destacar que a migração das empresas e seus respectivos produtos digitais para uma estrutura assente em base de dados ainda está em processo. O ritmo de adoção e transferência é gradual. Mais acelerado para empresas com histórico de investimento em pesquisa e desenvolvimento, isto é, os grandes conglomerados e grupos de mídia, assim como para empresas totalmente digitais, ocupando posição privilegiada a partir da criação de produtos inovadores que se tornaram referência conceitual e tecnológica.

Além delas, há as experiências de ponta, idealizadas em âmbito acadêmico ou por pesquisadores individuais e que são igualmente inovadoras. As empresas informativas de médio e pequeno portes, talvez, sejam as que realizem a migração de maneira mais lenta, embora possam existir exceções. De um modo ou de outro, mais cedo ou mais tarde, a tendência natural é que a conformação desse padrão seja efetivada nas organizações jornalísticas, por se tratar de uma decisão estratégica.

Antes de passarmos ao tópico com a conceituação operacional de base de dados e a explicação do modelo **JDBD**, é importante ressaltar que a definição de base de dados que colocamos em operação tenta ser mais abrangente. E, para isso, como ponto de partida, nos baseamos nas concepções sobre essa tecnologia da informação tal como ela é tratada na literatura das ciências da computação (DATE, 1984; FRANGOMENI, 1987; PEREIRA, 1998; SOL, 1998; MAURER; SCHERBAKOV; HALIM, RAZAK, 1998; SOUSA, 2002; GUIMARÃES, 2003; PIATTINI; CALERO, 1999; PIATTINI, 2000; BRICKLIN, 2001; VILLEGAS Y GARCÍA, 2002; CARLEY, 2002; JORGENSEN, 2003; SALGADO, RODRÍGUEZ, 2003; PAUL, 2004; GRAY, 2004).

Para chegarmos à sua formulação, aliamos aquelas concepções ao aporte teórico provido pelos trabalhos de pesquisadores que têm estudado a aplicação desta tecnologia

no jornalismo digital de modo mais direto ou indireto (COLLE, 2002, 2005a, 2005b; PALACIOS, 2002a, 2003b, 2003c, 2004; MACHADO, 2004a, 2004b, 2004c, 2006; FIDALGO, 2003, 2004; LÓPEZ, GAGO, PEREIRA, 2003; QUADROS, 2004, 2005; LIMA JÚNIOR, 2004, 2005, 2006; GARCÍA et al, 2005; HOLOVATY, 2005, 2006a, 2006b; GAGO, 2006). Agregamos, ainda, referências de outros investigadores da área da comunicação e do jornalismo (KOCH, 1991; CARITÁ, 1994; PAUL, 1999; GARRISON, 1998; HERBERT, 2000; REAVY, 2001; PAVLIK, 2001, 2005; LAGE, 2001; 2002; QUINN, 2002, 2005; GUNTER, 2003; PRYOR, 2003; QUIJADA, 2006), assim como das ciências da informação (SEMONCHE, 1993a, 1993b; RECODER; ABADAL; CODINA, 1995; PUJOL, 1995; SELLARÉS, 1995; MARCOS RECIO, 1999; TRAMULLAS Y KRONOS, 1997, 2000).

#### **4.1.1 A definição e a concepção operacional de “JDBD”**

Para o jornalismo digital, as bases de dados são definidoras da estrutura e da organização, bem como da apresentação dos conteúdos de natureza jornalística. Elas são o elemento fundamental na constituição de sistemas complexos para a criação, manutenção, atualização, disponibilização e circulação de produtos jornalísticos digitais dinâmicos.

Como tais, desempenham um conjunto de funções percebidas tanto quanto à gestão interna dos produtos, quanto aos processos de apuração e contextualização, à estruturação das informações, à composição das peças informativas, assim como à recuperação das informações e à apresentação dos conteúdos. Quais sejam:

- 1 - Indexar e classificar as peças informativas e os objetos multimídia;
- 2 - Integrar os processos de apuração, composição e edição dos conteúdos;
- 3 - Conformar padrões novos para a construção das peças informativas;
- 4 - Agilizar a produção de conteúdos, em particular os de tipo multimídia;
- 5 - Propiciar categorias diferenciadas para a classificação externa dos conteúdos;
- 6 - Estocar o material produzido e preservar os arquivos (memória), assegurando o processo de recuperação das informações;
- 7 - Permitir usos e concepções diferenciadas para o material de arquivo;
- 8 - Garantir a flexibilidade combinatória e o relacionamento entre os conteúdos;
- 9 - Gerar resumos de notícias estruturados e/ou matérias de modo automatizado;
- 10 - Armazenar anotações semânticas sobre os conteúdos inseridos;
- 11 - Habilitar o uso de metadados para análise de informações e extração de conhecimento, seja por meio de técnicas estatísticas ou métodos de visualização e exploração como o *data mining*;
- 12 - Ordenar e qualificar os colaboradores e “repórteres cidadãos”;
- 13 - Orientar e apoiar o processo de apuração, coleta, e contextualização dos conteúdos;
- 14 - Regular o sistema de categorização de fontes jornalísticas;
- 15 - Sistematizar a identificação dos profissionais da redação;
- 16 - Cartografar o perfil dos usuários;
- 17 - Transmitir e gerar informação para dispositivos móveis (celulares, computadores de mão, *iPods*, entre outros);

## 18 - Implementar publicidade dirigida.

Para a identificação das funções apontadas acima, aplicamos um esquema articulado entre a observação empírica dos 22 produtos jornalísticos digitais de diferentes perfis que integram o conjunto de casos agregados na amostra trabalhada na tese e o amparo instrumental na abordagem teórica desenvolvida. Um processo continuamente retroalimentado pelo estudo teórico e pela verificação da prática, bem como das experiências em curso.

Vale notar que as funcionalidades das bases de dados para o emprego no jornalismo digital estão afinadas com o que alguns especialistas das ciências da computação denominam como a terceira geração na linha evolutiva desta tecnologia da informação (PIATTINI, 2000; SALGADO, RODRÍGUEZ, 2003). As bases de dados de terceira geração se caracterizam por proporcionar capacidades de gestão de dados, objetos, gestão do conhecimento e atendem às necessidades de aplicações para informação textual, publicação digital, sistemas estatísticos, sistemas de informação geográfica, educação e formação, comércio eletrônico, entre outras.

Os avanços são reconhecidos em torno de três dimensões, conforme Mário Piattini (2000, p.01):

- Dimensão de rendimento: crescimento exponencial, aliado aos avanços de *hardware* e seu barateamento;
- Dimensão de inteligência: aumento das funcionalidades, associada à integração semântica dos dados, ao aspecto da segurança, à consistência das informações e ao fator tempo; e
- Dimensão de distribuição: a internet e a web definem novos conceitos, como a mobilidade e a descentralização.

Quanto à classificação das bases de dados no jornalismo digital – conforme especificações técnicas explicadas no subtópico “3.1.4 Tipos de bases de dados” –, entendemos que elas são: ativas, dedutivas, distribuídas, móveis, relacionais, multimídias referenciais e descritivas, e dinâmicas (PIATTINI; CALERO, 1999; PIATTINI, 2000; TRAMULLAS Y KRONOS, 1997, 2000; VILLEGAS Y GARCÍA, 2002; SALGADO, RODRÍGUEZ, 2003; GUIMARÃES, 2003; JORGENSEN, 2003; QUIJADA, 2006).

A delimitação das funções foi pensada e operacionalizada a partir de algumas noções que, inicialmente, propomos como características do **Jornalismo Digital em Base de Dados**: resolução semântica, metadados, relato imersivo ou narrativa multimídia, e jornalismo participativo (BARBOSA, 2005a, 2005b, 2005c, 2006a, 2006b). Também, procuramos integrar as três funcionalidades apontadas por Machado (2004a, 2004b, 2006): de formato para a estruturação da informação; de suporte para modelos de narrativa multimídia; e de memória dos conteúdos publicados. Além dessas, incorporamos, de certa maneira, aquela que foi trabalhada por Luciana Moherdau (2005): de mapear os usuários.

Da associação estabelecida entre as noções características iniciais - em especial a de resolução semântica - e o conjunto de funcionalidades que apresentamos, encontramos sete categorias descritivas, que, para nós, são expressivas do modelo **Jornalismo Digital em Base de Dados (JDBD)** e alcançam objetividade dado o seu caráter operativo. Estas categorias são:

- Dinamicidade;
- Automatização;
- Inter-relacionamento/Hiperlinkagem;

- Flexibilidade;
- Densidade informativa;
- Diversidade temática;
- Visualização.

Assim, temos como conceitualização para o **Jornalismo Digital em Base de Dados (JDBD)** a que segue:

- O **JDBD** é o modelo que tem as bases de dados como definidoras da estrutura e da organização, bem como da apresentação dos conteúdos de natureza jornalística, de acordo com funcionalidades e categorias específicas, que vão permitir a criação, a manutenção, a atualização, a disponibilização e a circulação de produtos jornalísticos digitais dinâmicos.

Trata-se de um paradigma na transição entre a terceira e a quarta geração do jornalismo digital. Como tal, delineia, também, uma metáfora particular que está relacionada com os modos diferenciados para a visualização das informações.

Elias Machado (2006), em seu recente livro, definiu o que ele também denomina “Jornalismo Digital em Base de Dados” como sendo a modalidade jornalística que utiliza as Bases de Dados para estruturar e organizar todas as etapas do processo de produção jornalística: apuração, composição, edição, circulação (MACHADO, 2006, p.08).

Tanto esta definição como a que propomos ressaltam a especificidade das BDs para o jornalismo digital, enfatizando esta tecnologia como um aspecto-chave, ao mesmo tempo em que contemplam uma percepção mais ampla na sua significação. O

que, para nós, reitera o potencial de reinvenção no uso desta tecnologia no jornalismo, confirmando o movimento ascendente e contínuo da curva do “S”, conforme argumentamos.

O quadro a seguir ilustra melhor o conceito, as funcionalidades e as respectivas categorias do modelo **Jornalismo Digital em Base de Dados (JDBD)**, que serão descritas nos tópicos subseqüentes:

<b>JORNALISMO DIGITAL EM BASE DE DADOS (JDBD)</b>	
<b>Conceito Operacional</b>	Modelo que tem as bases de dados como definidoras da estrutura e da organização, bem como da apresentação dos conteúdos de natureza jornalística, de acordo com funcionalidades e categorias específicas, que vão permitir a criação, a manutenção, a atualização, a disponibilização e a circulação de produtos jornalísticos digitais dinâmicos.
<b>FUNCIONALIDADES</b>	<b>CATEGORIAS</b>
<b>Indexar e classificar as peças informativas e os objetos multimídia;</b> <b>Integrar os processos de apuração, composição e edição dos conteúdos;</b> <b>Conformar padrões novos para a construção das peças informativas;</b> <b>Agilizar a produção de conteúdos, em particular os de tipo multimídia;</b> <b>Propiciar categorias diferenciadas para a classificação externa dos conteúdos;</b> <b>Estocar o material produzido e preservar os arquivos (memória), assegurando o processo de recuperação das informações;</b> <b>Permitir usos e concepções diferenciadas para o material de arquivo;</b> <b>Garantir a flexibilidade combinatória e o relacionamento entre os conteúdos;</b> <b>Gerar resumos de notícias estruturados e/ou matérias de modo automatizado;</b> <b>Armazenar anotações semânticas sobre os conteúdos inseridos;</b> <b>Habilitar o uso de metadados para análise de informações e extração de conhecimento, seja por meio de técnicas estatísticas ou métodos de visualização e exploração como o <i>data mining</i>;</b> <b>Ordenar e qualificar os colaboradores e “repórteres cidadãos”;</b> <b>Orientar e apoiar o processo de apuração, coleta, e contextualização dos conteúdos;</b> <b>Regular o sistema de categorização de fontes jornalísticas;</b> <b>Sistematizar a identificação dos profissionais da redação;</b> <b>Cartografar o perfil dos usuários;</b> <b>Transmitir e gerar informação para dispositivos móveis, como celulares, computadores de mão, iPods, entre outros;</b> <b>Implementar publicidade dirigida.</b>	<b>Dinamicidade</b> <b>Automatização</b> <b>Inter-relacionamento/Hiperlinkagem</b> <b>Flexibilidade</b> <b>Densidade informativa</b> <b>Diversidade temática</b> <b>Visualização</b>

Quadro 6 – JDBD: Conceito, Funcionalidades e Categorias

## 4.2 AS FUNCIONALIDADES

Como já mencionado acima, as funcionalidades foram identificadas a partir dos trabalhos de investigadores que têm estudado as bases de dados e sua aplicação efetiva no jornalismo digital, enquadrando as BDs tanto como tecnologia como na sua noção de ferramenta, que agrega valor incomensurável ao processo de produção e recuperação das informações jornalísticas, na justa medida em que também desempenha papel importante para a oferta de conteúdos aos usuários.

Desta forma, incorporando o aporte teórico que fundamenta a construção conceitual que embasa a nossa investigação - associada à observação dos produtos - foi possível reconhecer 18 funcionalidades, relacionadas aos níveis de pré-produção, produção, disponibilização/circulação, apresentação, consumo e pós-produção (que corresponde à repercussão que os conteúdos alcançam<sup>222</sup> na blogosfera, e em *sites* sociais, como o *Digg*, por exemplo). Certamente, outros estudos poderão vir a identificar outras ou até mesmo percebê-las de modo mais condensado. No entanto, para os objetivos desta tese, consideramos que elas correspondem ao que encontramos na bibliografia estudada e está de acordo com o que identificamos no *corpus* empírico.

Algumas delas, provavelmente, em alguns dos produtos analisados, ainda podem estar no nível de possibilidades ou potencialidades. Ademais, consideramos que tais funcionalidades auxiliam na melhor percepção de como as bases de dados servem ao jornalismo, de modo geral, e ao jornalismo digital, de modo particular. Cabe ainda

---

<sup>222</sup> Por meio do uso do *TrackBack*, um tipo de *LinkBack*, que permite que se receba notificação quando alguém coloca um *link* para determinado artigo ou outro material contido no *site*. Também assegura à empresa proprietária de um *site* ou autor de um *blog* saber quem está linkando ou remetendo para seu conteúdo.

salientar que a seriação para a apresentação das funcionalidades não configura uma hierarquização, até porque muitas delas estão contidas em diversos níveis, perpassando todo o processo.

#### 4.2.1 Descrevendo as 18 funções identificadas

##### 1 - Indexar e classificar as peças informativas e os objetos multimídia:

Etapa essencial, pois será determinante para todo o trabalho de estruturação e organização, apresentação e recuperação do material jornalístico. Aqui, define-se como será feita a classificação dos registros correspondentes ao material jornalístico e aos respectivos complementos multimídia (fotografias, áudios de entrevistas, vídeos, *slide shows*, infografias estáticas ou interativas, mapas, entre outros). A maneira como a indexação e a classificação são feitas também ajudará no estabelecimento das relações e cruzamentos entre os registros inseridos, isto é, entre as matérias, reportagens, e os respectivos objetos multimídia. Além disso, terá correspondência direta na maneira como as informações serão apresentadas, no processo de classificação externa. No *Topix.net*, agregador de notícias funcionando de modo totalmente automatizado, podem-se ver, na seção “Browse All Topix”, todos os assuntos indexados pelos algoritmos e robôs “aranhas” (*crawlers*). Eles estão colocados sob o título “News Channel Guide” e mostram a imensa variedade de possibilidades para se inserir registros referentes a uma notícia.

## **2 - Integrar os processos de apuração, composição e edição dos conteúdos:**

Ao incorporar as bases de dados na gestão dos produtos jornalísticos digitais, as empresas têm a possibilidade de integrar as rotinas, da produção da pauta, à apuração, à elaboração dos textos, até a edição do material. Para isso, é necessário empregar BDs diferenciadas e, claro, ter jornalistas e profissionais da área de comunicação integrados às equipes de desenvolvimento dos sistemas de gestão que darão sustentação ao gerenciamento dos processos.

## **3 - Conformar padrões novos para a construção das peças informativas:**

Tal funcionalidade corresponde ao emprego das BDs como suporte para garantir a produção de narrativas interativas – a hipernarrativa - de maneira mais afinada com os recursos proporcionados pelo hipertexto, pelo espaço navegável, pelo suporte digital, enfim. Aqui também devem ser observadas as propriedades do jornalismo digital, que, como vimos, podem já estar incorporadas nos sistemas de gestão de conteúdos, a partir dos *templates* criados.

## **4 - Agilizar a produção de conteúdos, em particular os de tipo multimídia:**

Esta função está diretamente ligada à anterior, porém a destacamos com a intenção de enfatizar a importância da informação estruturada, armazenada em BDs, pois permitirá a automatização dos processos de produção de conteúdos, conferindo maior agilidade ao trabalho jornalístico. Cabe notar que essa utilização pode ocorrer em produtos de perfis diversos, estejam eles operando com um nível baixo, parcial, ou total

de automatização nas etapas de apuração, composição, edição ou apresentação dos conteúdos.

### **5 - Propiciar categorias diferenciadas para a classificação externa dos conteúdos:**

Quanto mais abrangente for a classificação para os registros (notícias e todo o conteúdo jornalístico) inseridos numa BD, maiores possibilidades se terá para a apresentação deles segundo elementos conceituais diversificados, para além dos tradicionais “Política”, “Economia”, “Esportes”, “Mundo/Internacional”, “Entretenimento”, “Ciência e Tecnologia”, por exemplo. Percebem-se algumas novas categorias criadas para organizar a oferta informativa, por exemplo, para o material de arquivo: “O que já foi notícia”, “O que fez esse dia entrar para a história”, “On this day”, “A fondo”; para os dados extraídos a partir das estatísticas de acesso, como as seções “Most popular”, “Dez mais”, “Les +”, “+ Lidos”, “Lo más visto”; para a produção dos usuários: “Jornalismo cidadão”, “Participa”, “vc repórter”, “Vós réactions”, “First Person”, entre outras, como: “Offbeat News Forum”, que lista notícias consideradas mais originais ou excêntricas; além daquelas que exploram *clusters* de notícias - como no agregador *News Explorer* do EMM -, oferecendo análise das notícias diárias, em diversos idiomas e em diferentes períodos de tempo. Ele permite ver as relações que se estabelecem entre personalidades, organizações, além de citações mais destacadas filtradas a partir do material jornalístico que agrega; as análises de acontecimentos através de linhas do tempo ou *time lines* que permitem ver as notícias por continente e país, inclusive o país mais referenciado nas notícias em determinado momento do dia. Tem-se, ainda, “Recommended stories”, seção que filtra os artigos mais lidos, conforme também recomendam os usuários, como apresentado pelo

*MSNBC*. Além disso, há as infografias interativas, os *blogs*, áudios, vídeos, *slideshows*, *podcastings*, *RSS*, alertas de notícias, entre outras.

**6 - Estocar o material produzido e preservar os arquivos (memória), assegurando o processo de recuperação das informações:**

Essa é uma função-chave para todas as empresas informativas. São elas as grandes produtoras de conteúdo, e, por isso, a digitalização e a organização de todo material de arquivo em BDs cumprem não apenas aos propósitos da organização em si, mas têm o importante papel de atualização da memória social, pois é às empresas informativas que recorrem os usuários para realizar pesquisas, nos mais diversos níveis. Como vimos, as BDs geram conhecimento e abrem novas possibilidades para a integração dos usuários de diversos setores da sociedade com os meios. Além disso, o arquivo assegura, por outro lado, incremento de receita para as operações digitais, por meio da venda de artigos ou do acesso para se pesquisar e ler o material de edições passadas, além de outras formas de comercialização possíveis.

**7 - Permitir usos e concepções diferenciadas para o material de arquivo:**

Uma vez digitalizado, armazenado, organizado e estruturado em BDs, o material de arquivo fica apto para ser trabalhado de maneiras diversas. Boa parte dos produtos analisados no *corpus* empírico desta tese, como *The New York Times*, *Washingtonpost.com*, *BBC News*, *Elpaís.com*, *Elmundo.es*, *Portal Terra*, *Folha Online*, *International Herald Tribune*, *News Brief*, *News Explorer*, *GoogleNews*, *Topix.net*, *Newsmap*, e *Libération*, *Blog Deu no Jornal*, fornece concepções diferenciadas para o material de arquivo, seja por meio de novas seções, seja por meio de *links* de aprofundamento oferecidos junto aos artigos, a análise dos acontecimentos em

determinados períodos de tempo, ou a produção de material analítico a partir dos principais acontecimentos do dia ou da semana, entre outros. Algumas seções são denominadas como: “O que fez esse dia entrar para a história”, “O que já foi notícia”, “On this day”, “A fondo”, etc. O material de arquivo oferece contexto para as matérias e reportagens, aparecendo integrado a elas a partir da inserção de *links* relacionados para o material de um produto específico e também para o dos concorrentes (como fazem *Elpaís.com* e *Elmundo*).

## **8 - Garantir a flexibilidade combinatória e o relacionamento entre os conteúdos**

O uso de BDs relacionais, associadas a outros tipos, proporciona grande capacidade para se realizar cruzamentos e combinações diversas entre as informações de natureza jornalística. O potencial de interrelacionamento e remetimentos é incomensurável e impulsiona a geração de conteúdos originais. Pode-se, por exemplo, oferecer os assuntos relacionados a uma determinada notícia; os *links* de aprofundamento; mostrar as matérias e reportagens que estão em desenvolvimento; saber quais são os artigos mais recomendados; os mais lidos; os áudios mais escutados; os vídeos, os *slideshows*, os infográficos mais vistos; as matérias que estão sendo enviadas por *e-mail*; saber quais são os artigos mais postados em *sites* como o *Digg* e nos *blogs*, conferindo a posição de relevância deles no *ranking* do *Technorati*<sup>223</sup>; saber quais outros artigos estão sendo lidos por quem também está lendo uma determinada matéria ou reportagem num certo momento; ler matérias relacionadas em outros *sites*

---

<sup>223</sup> <http://technorati.com/search/www.neuralgourmet.com>. O *Technorati* é uma espécie de mecanismo de busca que retrata o estado da *blogosfera*, além de outros *sites* independentes. Ele organiza os *blogs* e apresenta as estatísticas referentes a eles, como o número de *blogs* existentes, o número de *posts* que são publicados por dia, a frequência das atualizações, o *ranking* dos *blogs* mais populares, os assuntos e personalidades mais pesquisados nos *blogs*, etc.

noticiosos e em outros idiomas; ver as relações que são estabelecidas em torno de personalidades, como presidentes, ministros, empresários, dirigentes, organizações, entre outras possibilidades.

#### **9 - Gerar resumos de notícias estruturados e/ou matérias de modo automatizado:**

Funcionalidade mais amplamente percebida em produtos noticiosos que funcionam de maneira totalmente automatizada. São os chamados agregadores de notícias, que, empregando agentes inteligentes, isto é, algoritmos, para vasculhar a web a partir de parâmetros definidos, podem indexar os resultados segundo as categorias temáticas predefinidas (casos do *Google News*; *News Explorer*, *News Brief*, *Newzbubble*, *Newsmap* e, num nível mais básico, também *Notibits*) ou de acordo com um perfil temático específico (conteúdo local, como faz o *Topix.net*). Em geral, os agregadores funcionam coletando informações de grandes conglomerados da mídia e de sites noticiosos de referência (*Associated Press*, *Reuters*; *BBC News*; *CNN*; *Guardian Unlimited*; *Elpaís.com*; *Folha Online*; *Estadão*, entre muitos outros. Já o *Newsmap* é um agregador que coleta informações provenientes de outro agregador, o *GoogleNews*). Existem, por outro lado, casos em que algumas empresas, como a *Thomson Financial* e a própria *Reuters*, utilizam as suas respectivas BDs para gerar artigos, matérias, entre outros.

#### **10 - Armazenar anotações semânticas sobre os conteúdos inseridos:**

Para cada registro - informação ou objeto multimídia – inserido em uma BD armazenam-se, em paralelo, os dados que descrevem aquela informação ou objeto: o nome do autor, ou do produtor, a origem, ou mesmo a descrição do próprio conteúdo

em si da informação ou objeto multimídia. São os metadados ou metainformações (a matéria-prima das BDs) que descrevem e permitem categorizar as informações e, assim, possibilitam que sejam efetuadas a organização semântica, a partir de determinados padrões, e também a formatação dos dados, ou seja, a parte sintática.

**11 - Habilitar o uso de metadados para análise de informações e extração de conhecimento, seja por meio de técnicas estatísticas ou métodos de visualização e exploração como o *data mining*:**

Como matéria-prima das BDs, os metadados ou metainformações oferecem um grande potencial para ser aplicados aos conteúdos jornalísticos para a descoberta de relações e de conhecimento que podem estar escondidos por entre as informações. Para a extração de conhecimento a partir dos registros inseridos nas BDs, também são aplicadas técnicas de inteligência artificial, como as ontologias<sup>224</sup>, entre outras, que podem revelar informações valiosas. As técnicas de *data mining*, como temos visto, são aplicadas tanto para garimpar informações escondidas na parte estrutural das BDs, como para representar tais dados em padrões diversos de visualização, conformando o que se denomina “estética base de dados”. Também o jornalismo pode e deve tirar partido dessa estética para apresentar as informações segundo novas matrizes, como tem feito alguns produtos que constam da nossa amostra. Há, ainda, outras técnicas de visualização de informações como a *treemap*, empregada pelo site *Newsmap*. Consideramos que tais técnicas auxiliam no estabelecimento de uma nova metáfora para os produtos jornalísticos digitais.

---

<sup>224</sup> Conhecimento sobre relações entre palavras, dados, objetos. Uma ontologia é um vocabulário controlado que descreve objetos e as relações entre eles em um modo formal. São especificações de conceitualizações utilizadas para o compartilhamento do conhecimento. In: *WhatIs?com* ([http://whatis.techtarget.com/definition/0,,sid9\\_gci212702,00.html](http://whatis.techtarget.com/definition/0,,sid9_gci212702,00.html)).

## **12 - Ordenar e qualificar os colaboradores e “repórteres cidadãos”:**

Seguindo a proposta de Maria Angeles Cabrera (2005), pensamos que é possível ordenar não apenas o nível de participação que aos usuários possa ser facultado por um produto jornalístico digital, mas, estendendo essa possibilidade, percebemos que as BDs podem ser empregadas para ordenar e qualificar também aqueles colaboradores (colunistas, articulistas, etc.) habituais ou eventuais que um jornal, revista, portal ou outro tipo de *site* jornalístico possa ter ou querer ter em seus quadros.

## **13 - Orientar e apoiar o processo de apuração, coleta, e contextualização dos conteúdos:**

Significa empregar as BDs como fontes para a realização do trabalho jornalístico. Embora as técnicas da *CAR* tenham sido desde há muito incorporadas ao processo de produção no jornalismo, com a expansão das BDs e dos conteúdos nelas armazenados e disponibilizados, é essencial que os profissionais as integre ainda mais em seu dia-a-dia, para a pesquisa, análise e contextualização das informações. Isso vale não apenas para os profissionais de produtos digitais, mas para todos os jornalistas.

## **14 - Regular o sistema de categorização de fontes jornalísticas:**

Se cada jornalista inserir as informações correspondentes às fontes que utiliza, sejam elas oficiais, oficiosas ou independentes, o produto poderá regular o sistema de categorização delas, classificando-as por áreas, tipo, especialidade, etc.; assim como a inserção dos respectivos números de telefone, *e-mail*, *web sites*, entre outros dados que possam ajudar na melhor identificação das fontes e, deste modo, contribuir para que se

tenha um melhor ordenamento do trabalho de apuração e a garantia de maior agilidade ao processo de produção jornalístico.

### **15 - Sistematizar a identificação dos profissionais da redação:**

Com o uso de BDs, o processo de identificação de todos os profissionais de uma redação fica facilitado, pois permite organizar melhor as informações sobre todo o corpo de repórteres nos mais diferentes níveis e funções, assim como editores, coordenadores, diretores, etc. Os dados armazenados e organizados asseguram, inclusive, a adequação dos processos de atribuição de autoria do material, com maior precisão, controle e confiabilidade. Traz, ainda, benefícios quando, por exemplo, se for vender algum conteúdo para alguma outra empresa, garantindo a preservação dos direitos autorais.

### **16 - Cartografar o perfil dos usuários:**

A partir do registro e cadastramento dos usuários nos produtos, com dados básicos, como nome, cidade em que mora, assuntos do seu interesse, entre outros, além daqueles dados que são colhidos quando eles navegam pelas seções de um *site* jornalístico, é possível criar uma considerável BD e, a partir dela, cartografar o perfil deles. Com isso, podem-se determinar perfis dinâmicos, como: criar classes de usuários, com base na ordem de leitura das seções; preparar conteúdos e serviços dirigidos a tais grupos, entre outros. No *Topix.net*, por exemplo, os algoritmos realizam a categorização dos assuntos de modo multifacetado, fazendo com que os conteúdos possam ser disponibilizados para usuários de áreas, países e cidades distintas. Essa personalização é imediata, isto é, o *site* identifica instantaneamente de onde o usuário está acessando e

lhe mostra assuntos relacionados ao seu país, e à sua cidade, além dos assuntos mais gerais.

**17- Transmitir e gerar informação para dispositivos móveis (celulares, computadores de mão, *iPods*, entre outros):**

Com as informações estruturadas, organizadas e mantidas nas BDs, a disponibilização e a geração de informação personalizada para outros dispositivos são facilitadas. Como vimos, tanto a partir de Manovich (2001) e mesmo com a descrição das vantagens de recursos possibilitados por linguagens de programação como a *XML*, se pode separar o conteúdo da forma e, assim, criar produtos diferenciados para uma gama variada de dispositivos. Para cada um deles, uma interface diferente será desenhada para os conteúdos que estão estruturados nas BDs. Note-se que tais conteúdos podem ser disponibilizados a partir de perfis identificáveis sobre as características de cada uma das versões personalizadas pretendidas. Em geral, a maioria dos casos observados fornece conteúdos para plataformas móveis.

**18 - Implementar publicidade dirigida:**

A partir da cartografia do perfil dos usuários, pode-se direcionar, personalizar, o conteúdo publicitário e, assim, alcançar maior efetividade. Ou seja, sabendo-se previamente qual grupo de usuários tem mais interesse em determinado tipo de produto, ou de onde esses usuários estão acessando, a estratégia de marketing pode se tornar mais eficaz, assim como a canalização de anunciantes para o *site*. Quanto mais conhecerem o perfil dos seus usuários, maiores chances as empresas informativas terão de implementar a publicidade e de atrair anunciantes para os respectivos produtos. Boa

parte dos *sites* tem anúncios gerados pelo serviço de anúncios do *Google*, o que lhes permite endereçar *links* publicitários dirigidos. No *Topix.net*, observamos que há a personalização da publicidade, assim como de classificados.

### 4.3 AS CATEGORIAS

Para chegarmos às sete categorias propostas como específicas do **Jornalismo Digital em Base de Dados (JDBD)**, procedemos, inicialmente, ao mapeamento de quatro características (resolução semântica, metadados, relato imersivo ou narrativa multimídia e jornalismo participativo), que nos permitiram, em seguida, aplicá-las como modelos de apreensão do objeto em estudo. Apesar do valor heurístico nelas contido, constatamos que precisaríamos identificar categorias mais expressivas, que pudessem alcançar objetividade e, com isso, nos fornecer elementos para melhor analisarmos os produtos.

Assim, utilizando aquele modelo de apreensão – no qual resolução semântica (FIDALGO, 2003, 2004) é a noção central – de maneira articulada com as funcionalidades e a abordagem teórica, chegamos às categorias. Entendemos que o estabelecimento de tais categorias demarca e complementa as particularidades do modelo **JDBD**. Além disso, são essenciais para atribuir e reconhecer as rupturas, as remediações ou continuidades, bem como as potencialidades geradas pelo uso de base de dados no jornalismo digital.

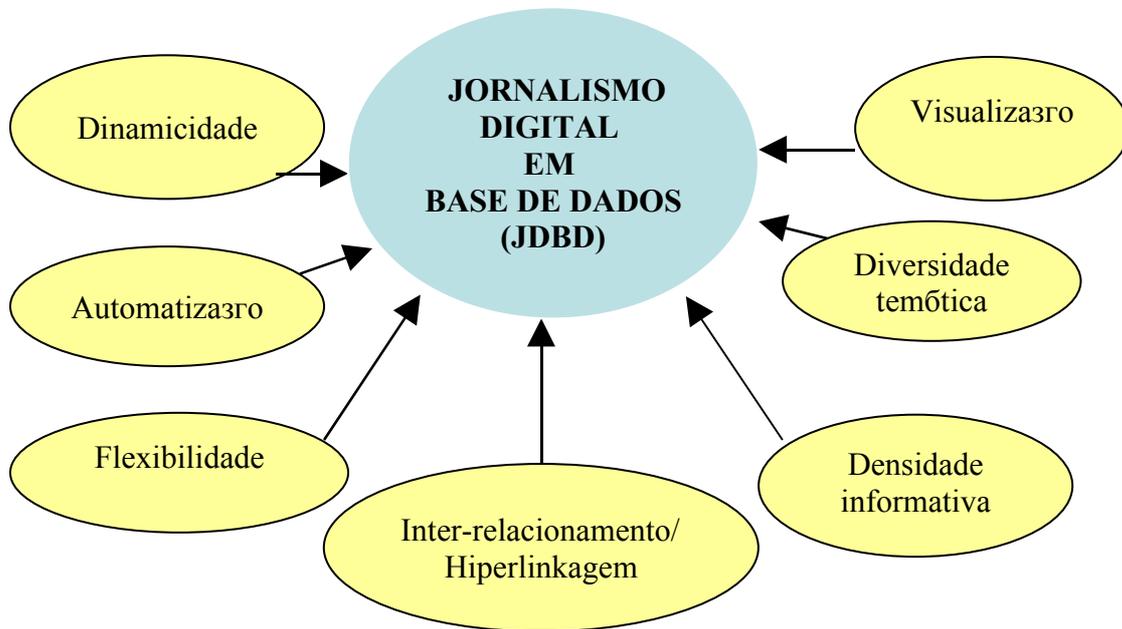


Figura 42 – Categorias do **Jornalismo Digital em Base de Dados (JDBD)**

A seguir, procedemos à apresentação de cada uma das sete categorias.

#### 4.3.1 Dinamicidade

Considerando-se o caráter dinâmico intrínseco às bases de dados, esta primeira categoria, de fato, guarda o aspecto diferencial para promover a adequação dos processos de implementação dos *web sites* e do produto jornalístico digital em si. É com a adoção das BDs que se pode obter maior automatização para os processos, através da integração de linguagens de programação para a construção dos sistemas de gestão de conteúdos, nos quais todas as especificações conceituais e técnicas estarão formatadas

para a produção das informações, redação, edição e apresentação do material. Conseqüentemente, se poderá trabalhar com a oferta de conteúdos mais originais e afinados com as propriedades do jornalismo nas redes digitais.

Como afirma Manuel Gago (2006, p.99), um *site* dinâmico deve flexibilizar a forma de trabalhar e agilizar todos aqueles processos mecânicos que não necessitem da capacidade intelectual e analítica humana. “A partir de esa armazón de automatismos, establecer una profunda estructura editorial que permita la narración de historias que aprovechen las características tecnológicas de la red”<sup>225</sup>. Complementando, cabe notar que *site* dinâmico significa conteúdo em renovação constante, isto é, oferta de mecanismos de busca refinados, inter-relacionamentos entre as informações, disponibilização de estatísticas ajustadas a cada instante à medida em que se ampliam os conteúdos da base de dados, as já estabelecidas *breaking news* (notícias atualizadas continuamente), entre outros aspectos já mencionados na descrição das funcionalidades, além de outros que veremos a seguir.

Desta maneira, temos dinamicidade como categoria mestre a guiar a verificação de um novo padrão, em contraposição com aquele que vigorou anteriormente, ou seja, o dos *sites* estáticos. Sendo assim, é a partir dela que as demais categorias adquirem representatividade. Dentre os casos que constituem a nosso *corpus* empírico, é possível reconhecer desde aqueles que operam com um grau elevado de dinamicidade até os que ainda estão num nível mais básico.

Consideramos que os agregadores de notícias (*Google News*, *News Brief*, *News Explorer*, *Topix.net*, *Newzbubble*, *Newsmap*) apresentam o nível mais elevado de dinamicidade, pelo fato de terem seus processos num estado de automatização total. *Notibits*, por outro lado, é um produto que ainda está num estágio baixo de

---

<sup>225</sup> “A partir dessa armação de automatismos, establecer una profunda estructura editorial que permita a narración de histórias que aproveitem as características tecnológicas da rede”. (t.A.).

dinamicidade, por não estar estruturado em BDs dinâmicas, ao contrário da sua antecessora *TDC*, que operava em conformidade com um padrão de dinamicidade.

*Folha Online* e o *Portal Terra* estão num nível médio, pois empregam parcialmente os recursos que podem alçá-los a um grau maior de dinamicidade. O produto laboratorial *Urbi et Orbi* passou a ser um *site* dinâmico após a reformulação efetuada no primeiro semestre de 2006, o que teve consequência direta na composição das peças informativas, no acréscimo de temáticas (inclusive com a incorporação da seção “Ensino superior”, que era o foco informativo do antecessor *Akademia*), de recursos multimídia e de conteúdos oferecidos, assim como na apresentação das informações. Já o *Panopticon*, pelas implementações que vêm sendo realizadas na *Plataforma de Publicação e Ensino Panopticon (PPEP)*, avança para um estágio de alto grau de dinamicidade. O *blog Deu no Jornal*, por seu próprio perfil, nos permite considerá-lo num nível médio de dinamicidade. Embora seja ele próprio uma base de dados com notícias sobre corrupção, o *blog* não conjuga a oferta de recursos considerados como importantes para a *blogosfera*, como por exemplo as opções de compartilhamento em *sites* sociais.

*BBC News*, *Elpaís.com*, *Elmundo.es*, *International Herald Tribune*, *Libération*, *MSNBC*, *The New York Times* e *Washingtonpost.com* possuem um nível alto de dinamicidade, ainda que se observem variações que apontam para a maior amplificação desse nível em alguns deles, nomeadamente nos *sites* *BBC News*, *Elpaís.com*, *Elmundo.es*, *The New York Times* e *Washingtonpost.com*. Cada um a seu modo particular, vem experimentando com novas temáticas; com a implementação do uso dos arquivos; com a composição das peças informativas mais de acordo com uma narrativa interativa e multimídia; com os inter-relacionamentos entre os conteúdos; com a inserção do material produzido pelos usuários; com a oferta de conteúdos para

plataformas móveis, entre outros recursos agregados, a exemplo das opções de compartilhamento das informações em *sites* sociais. Esses produtos, talvez, sejam os representantes do *mainstream* da mídia que mais têm buscado inovar.

#### 4.3.2 Automatização

Inerente à incorporação das BDs aos processos de armazenamento, estruturação, organização e apresentação das informações e ligada diretamente à noção de dinamicidade. Percebemos três tipos de automatização: - automatização parcial, quando aplicada apenas a algumas etapas do processo de produção jornalístico (como para a apresentação dos conteúdos); - automatização procedimental (nível intermediário), na qual se trabalha com mais etapas do processo jornalístico de modo automatizado; e - automatização total, quando os produtos funcionam totalmente de maneira automatizada, sem qualquer intervenção dos jornalistas. Tem-se aqui o emprego dos metadados ou metainformações, do *data mining*, dentre outras técnicas para trazer à tona informações que ficam invisíveis e têm o potencial para conformar tematizações diversas para os conteúdos, além de modos mais originais e inovadores para apresentá-los. Cabe ressaltar que os metadados, assim como o *data mining*, entre outras técnicas, poderão estar integradas também nos outros tipos de automatização, a depender das definições conceituais e editoriais dos produtos. Note-se, ainda, que a automatização dos processos de produção no jornalismo resulta em ganho de tempo para que os jornalistas se dediquem ao trabalho intelectual e analítico.

### 4.3.3 Flexibilidade

A maleabilidade da tecnologia de bases de dados tem feito com que os desenvolvimentos expandam a sua adoção, nos mais diversos setores. No jornalismo, desde os anos 70, elas têm sido empregadas de modos diversos, assegurando maior agilidade, qualidade e flexibilidade para a produção dos conteúdos, bem como para a sua disponibilização. A capacidade de reinvenção para o emprego de BDs no jornalismo assegura-lhes a condição de aspecto-chave para a modalidade do jornalismo digital, na fase que designamos como intermediária entre a terceira e a quarta geração. Neste estágio, elas deixam de ser vistas apenas como soluções para o armazenamento de informações, e passam a ser definidoras da estrutura, da organização, bem como da apresentação dos conteúdos de natureza jornalística para viabilizar a criação, a manutenção, a atualização, a disponibilização e a circulação de produtos jornalísticos digitais dinâmicos. A flexibilidade também está relacionada a sistemas de apuração menos hierarquizados, à facilidade de recuperação dos conteúdos para a contextualização, a uma maior autonomia do trabalho dos profissionais, assim como à descentralização da produção, pois, com sistemas de gestão de conteúdos ancorados em BDs, pode-se publicar de qualquer lugar.

#### 4.3.4 Inter-relacionamento/Hiperlinkagem

Um dos grandes potenciais identificados para as BDs no jornalismo digital. A partir dos campos de classificação interna, tem-se a capacidade de identificar padrões combinatórios e inter-relacionamentos diversos entre as informações. A hiperlinkagem também colabora para a incorporação do material de arquivo à oferta informativa, permitindo maior aprofundamento e contextualização. Abrange, ainda, os elementos da hipernarrativa para a construção das peças informativas, assim como os recursos que são agregados (áudios, vídeos, fotos, *slideshows*, *infografia interativa*, *podcastings*, *RSS*, opções de compartilhamento das informações em *sites* sociais, *blogs*, e o material produzido pelos usuários em forma de comentários, textos, fotos, vídeos, dentre outros) para o relato dos acontecimentos e a sua apresentação. Em última instância, é por meio dos inter-relacionamentos entre os registros inseridos na etapa da indexação e classificação interna que se pode assegurar, paralelamente, tematizações novas para os conteúdos.

#### 4.3.5 Densidade informativa

Na base da definição desta categoria está o conceito de resolução semântica (FIDALGO, 2003, 2004). Ele é adequado para explicar a sucessão de notícias na informação *on-line* disponibilizada de maneira contínua. Nele, estão incluídas as noções de baixa e alta resoluções. Ou seja, uma notícia sobre um incêndio em um bairro de Salvador, por exemplo, será publicada, inicialmente, de modo imediato e deficiente, dando conta apenas de que houve um incêndio num determinado ponto da capital baiana. Porém, à medida que mais informações forem sendo apuradas, inclusive, pesquisando-se em BDs, ela vai sendo complementada, alterada, corrigida, contextualizada, aprofundada - tanto pelos profissionais de um produto, como pela participação dos usuários, com o envio de fotos e testemunhos, por exemplo - até de fato alcançar o nível em que passará da baixa resolução inicial para uma situação de alta resolução, quando se atinge um alto nível de consolidação, ou seja de densidade informativa. A densidade informativa também é assegurada por uma oferta informativa mais abrangente e diversificada.

Os produtos que, de fato, trabalham com uma diversidade maior de recursos associados à cobertura dos eventos, bem como ao tratamento, processamento e publicação do material, oferecem, naturalmente, uma densidade informativa superior. Entre os agregadores, consideramos que os produtos do *European Media Monitor* (*Newsbrief* e *NewsExplorer*) possuem uma variedade que os diferenciam e demonstram o seu caráter inovador. Enquanto o *NewsBrief* disponibiliza as *breaking news* atualizadas continuamente, e mostrando qual o nível de acesso para elas ao longo do dia, o *NewsExplorer* trabalha com a análise das notícias diárias através da organização

em *clusters* (“Countries”, “Related People”, “Related Organizations”, “Alerts”), explorando relacionamentos entre os personagens da notícia, bem como entre as organizações e países citados, e examina os acontecimentos por meio de linhas do tempo, e em diversos idiomas, entre outros.

Já o *Google News*, pela abrangência do material que indexa, proveniente de 4.500 publicações noticiosas, amplia o leque de possibilidades para se acompanhar a cobertura dos acontecimentos. O *Newsmap*, por outro lado, oferece a oportunidade para se comparar a cobertura feita por muitos *sites* jornalísticos de distintos países, através do emprego de técnica inovadora para se visualizar tais informações provenientes do *Google News*. O *Toipix.net*, por sua vez, traz como diferencial a atenção para o ordenamento dos conteúdos locais, personalizados de acordo com a região, país e cidade de onde o usuário está acessando, além de assuntos gerais. O *Newzbubble*, pelo nível elevado de dinamicidade, está continuamente renovando o seu conteúdo e gerando novas *tags* (marcas, etiquetas) conforme novas notícias ou desdobramentos dos acontecimentos vão sendo agregados no seu sistema. Essas *tags* são representadas pelo formato de “bolhinhas”, que serão maiores ou menores a depender da relevância do assunto.

O *site BBCNews*, em geral, trabalha com uma grande diversidade temática para explorar a cobertura dos acontecimentos, sempre agregando soluções e recursos inovadores para aproveitar o material gerado pela sua numerosa equipe, incorporando também o conteúdo produzido pelos usuários, além de experimentar com novos elementos para a organização da informação, bem como para a sua visualização.

No *Folha Online*, o *layout* e a maneira como o conteúdo é organizado na *home* ajudam na conformação da carga informativa. Se utiliza bastante o recurso do *link* para exibir as manchetes principal e secundárias. E, claro, para realizar o inter-

relacionamento entre as matérias e reportagens, sobretudo remetendo para o material de arquivo e para os especiais. O *Folha* trabalha com *blogs* dos seus próprios colunistas, além de mostrar as cinco matérias mais lidas (“+ lidas”), numa sinalização já de utilização de técnicas de extração de informações. No entanto, a oferta informativa poderia ser melhor implementada levando-se em conta o fato de ser um dos *sites* de referência no Brasil. O *Folha* ainda não possui espaço efetivo para as contribuições dos usuários (há apenas o “Painel do leitor”, onde se pode comentar assuntos previamente selecionados e o “Bate-papo”), nem tampouco incorpora material multimídia. As fotos são trabalhadas na “Galeria”, e ocasionalmente, gráficos estáticos são disponibilizados.

Já o *Portal Terra*, no que diz respeito ao conteúdo jornalístico, dá maior ênfase às últimas notícias, oferecendo possibilidade de contextualização para os assuntos, e trabalhando o material de arquivo nos especiais, além do subcanal “Fatos históricos”. Já integra as colaborações dos usuários (por meio da seção “vc repórter” aceita textos e fotografias), o que diversifica e aumenta a sua densidade informativa. Explora conteúdo em vídeo nos canais “Vídeo” e “Jornal do Terra”, além de fotografias, mas não oferece infografias. Os *blogs* estão integrados no canal “Terra Magazine”. Por seu turno, o *blog Deu no Jornal* apresenta o conteúdo, focalizado na temática “Corrupção”, organizado em quatro categorias: “Acesso a informações públicas”, “De olho na política”, “O melhor do jornalismo” e “Pra relaxar”, todas elas com análises que repercutem assuntos que foram noticiados e selecionados de 63 fontes jornalísticas na web, entre jornais e revistas. De acordo com o que se propõe, possui densidade regular, havendo potencial de expansão, a depender dos aprimoramentos.

Os espanhóis *Elmundo.es* e *Elpais* seguem um padrão de alta densidade informativa, pois, ambos, aplicam muito bem os recursos para a implementação da oferta informativa, seja por meio do inter-relacionamento entre os conteúdos, pelo uso

do material de arquivo, de *blogs*, pelo desenvolvimento de material multimídia, em especial as infografias interativas, uso do áudio, etc. O *Libération*, considerado um dos *sites* jornalísticos franceses mais inovadores, e sócio do *Elmundo*, vem implementando gradualmente o seu conteúdo, dando muito destaque aos *blogs*, às opções de interatividade para que os usuários participem ativamente, à contextualização dos assuntos tratados nas matérias, nas quais se trabalha com *links* de inter-relacionamento, e, ainda, à composição de material multimídia.

O *Washingtonpost.com* e o *The New York Times* são também exemplos de edições digitais de jornais com similares impressos que vêm investindo muito no aprimoramento dos respectivos conteúdos. Contam com valiosos arquivos e, em conformidade com seus modelos operacionais, possuem uma densidade informativa que podemos considerar superior em relação a alguns outros casos. A variedade da oferta informativa tem se ampliado com o material multimídia (*slideshows*, vídeos, áudios, etc) integrados à cobertura dos acontecimentos, à composição de narrativas interativas, além de experiências com aplicações “sociais”, e até com novelas (“Jezebel’s Tomb”<sup>226</sup>, caso do *Washingtonpost.com*), dentre outras mencionadas anteriormente nesta tese.

No *MSNBC* tem-se a inclusão de conteúdos destacados do *Washingtonpost.com*, da rede de TV *NBC* e da revista *Newsweek* ampliando a sua oferta que já é bastante variada, com peças informativas interativas, opções de compartilhamento, estatísticas dos conteúdos mais vistos e mais enviados por *e-mail*, dentre outras. Já no *International Herald Tribune*, a densidade informativa é menos consistente, embora se observe o uso dos *links* de contextualização nas peças informativas, como também de inter-

---

<sup>226</sup> <http://www.washingtonpost.com/wp-srv/onlinebooks/jezebelstomb/>. É a primeira vez que o *site* publica conteúdo ficcional. A novela, baseada em livro homônimo escrito pelo repórter da seção “Business” do jornal, David Hilzenrath, estreou no dia 17 de janeiro de 2007. As atualizações a partir das 412 páginas do livro são realizadas às segundas e às terças-feiras. Os usuários que quiserem adquirir a edição impressa de “Jezebel’s Tomb” têm que desembolsar U\$18.95.

relacionamentos, recursos para se ouvir os artigos, compartilhá-los, além da disponibilização de estatísticas dinâmicas.

#### **4.3.6 Diversidade temática**

Como a categoria anterior, esta também incorpora a noção de resolução semântica. Novas tematizações podem ser trabalhadas para assegurar também maior densidade informativa e vice-versa. Consideramos que tal categoria possui uma carga muito maior de potencialidade do que temos observado como real implementação na maioria dos produtos jornalísticos digitais analisados, os quais têm tentado ultrapassar o eixo temático mais comum utilizado para organizar os conteúdos, quais sejam: política, economia, esportes, cultura (FONTCUBERTA, 2006, p.72), além de ciência, saúde e tecnologia. Notamos que alguns *sites* já vêm recorrendo a elementos conceituais diversificados para dar forma a algumas seções ou espaços instituindo, desta maneira, novas temáticas. Isso evidencia que quanto maior for a incorporação de BDs para a estruturação do material jornalístico, para a construção das peças informativas e para a apresentação dos conteúdos, tanto maior serão as possibilidades para as tematizações.

### 4.3.7 Visualização

Diz respeito aos modos diferenciados para se representar informações jornalísticas, a partir da sua estruturação em bases de dados. A esta categoria são intrínsecas as noções de metadados ou metainformações e as técnicas de *data mining*, assim como outras que já vêm sendo empregadas, a exemplo da *tree map*, que gera o padrão de visualização conhecido como *Squarified* (retângulos alongados), tal como trabalhado pelo *Newsmap*. Já o *blog Deu no Jornal* trabalha com o *software* livre *LinkBrowser*<sup>227</sup> para gerar mapas de visualização que mostram os relacionamentos existentes entre as pessoas e os assuntos mencionados nas matérias contidas em sua base de dados.

Esta categoria ainda tem grande potencial de expansão, à medida que aqueles elementos forem mais abrangentemente empregados. No entanto, como mencionamos, alguns dos casos incorporados ao *corpus* empírico já dão uma amostra do que é possível realizar para sair do convencionalismo mais associado à metáfora do jornal impresso.

No capítulo a seguir, abordaremos mais detalhadamente a emergência do que, na nossa perspectiva, já delinea uma nova metáfora para os produtos jornalísticos digitais. Ao lado disso, examinaremos algumas rupturas e potencialidades que o modelo **JDBD** traz para o jornalismo digital.

---

<sup>227</sup> O aplicativo foi desenvolvido pela empresa TouchGraph (<http://sourceforge.net/projects/touchgraph>) e adaptada para uso - em ambiente Java - no *blog* pela empresa ECJ Tecnologia da Informação (<http://www.ecj.com.br/>).



## **UMA NOVA METÁFORA A TOMAR FORMA**

## 5

---

**Uma nova metáfora a tomar forma**

---

*The metaphor for the electronic presentation of news is still evolving and no one is certain what the most effective metaphor will be.*

- C.R. Watters; M.A. Sheperd; T. Chiasson; L. Manchester.

*(...) Rupturas não se resumem em apenas apresentar possibilidades até então inexistentes em outros suportes... é necessário que o webjornal modifique funções já existentes ou desenvolva novas funções.*

- Luciana Mielniczuk.

### 5.1 RUPTURAS QUALITATIVAS E POTENCIALIDADES DO “JDBD”

Já se passaram 12 anos desde que Mindy McAdams (1995) publicou, no *Interpersonal Computing and Technology: as electronic journal for the 21st century*, o artigo seminal *Inventing Online Newspaper* no qual recomendava a *newspaper metaphor* como o princípio fundamental a guiar a estruturação de *sites* jornalísticos. As razões apontadas para tal escolha incluíam a facilidade para usá-la e o fato de ser amigável e familiar, permitindo a qualquer pessoa reconhecer aquele padrão adotado nos jornais. A metáfora do impresso seria o guia para: organizar a disposição das chamadas na primeira página ou *front page* do *site*, destacando e apontando para os usuários os assuntos e pontos do sistema mais importantes; compartimentar os assuntos em seções, empregando inclusive denominações similares às já utilizadas pelos diários.

Apesar daquela recomendação ter tido como referência uma experiência específica (a do *The Washington Post*), o caminho apontado forneceu a receita que talvez muitos estivessem à espera para formatar seus produtos. Além da facilidade para se replicar uma fórmula já testada com sucesso, a também chamada *broadsheet metaphor* possuía elementos para se desenvolver interfaces que, para os usuários, seriam mais simples para obter informações jornalísticas através do computador.

Embora retrate um momento inicial, quando tudo parecia ainda nebuloso, o fato é que tal metáfora vingou como um padrão para os produtos jornalísticos digitais, como confirmam estudos posteriores e mais recentes (WATTERS et al, 1997; 2000; ERIKSEN & IHLSTRÖM, 2000; AKESSON, 2003; IHLSTRÖM; LUNDBERG, 2004). Passados 12 anos, com os aprimoramentos verificados – tanto com relação ao desenvolvimento de *softwares* entre outros aspectos de ordem tecnológica, quanto com a especialização dos profissionais – e a consolidação de uma nova modalidade jornalística no ciberespaço, acreditamos que o cenário atual, marcado pelo uso exponencial de bases de dados para estruturar, organizar e apresentar os conteúdos jornalísticos, indica a emergência de uma nova metáfora. A colaborar para o seu estabelecimento está o próprio nível de exigência dos usuários, que sempre esperam encontrar originalidade nos *sites* – seja nos conteúdos, seja nas opções de visualização para os mesmos.

Assim é que percebemos o **Jornalismo Digital em Base de Dados (JDBD)** com o potencial de também assegurar essa nova metáfora, fundamentada na *database aesthetics* (VESNA, 1999, 2000a, 2000b; LEÃO, 2003; PAUL, 2004; DIETZ, 2000, 2004). Isto é, um modo particular para a apresentação das informações jornalísticas já desvinculado da metáfora do impresso e que procede diretamente do emprego das BDs para a estruturação dos conteúdos no jornalismo digital. Certamente, ela não

determinará um padrão único, fácil de ser replicado, pois técnicas e métodos diversos podem ser utilizados para gerar visualizações diferenciadas, conforme pudemos constatar a partir da análise dos casos que integram o *corpus* empírico desta tese. Contudo, ela demarca uma estratégia comunicacional (GROSSMANN, 2006) mais apropriada à produção jornalística no ciberespaço, trazendo renovação, diversidade e contribuindo para o aprimoramento e a evolução de padrões visuais para a representação das informações.

Ao propormos o **JDBD** como um paradigma e apontar a emergência de uma nova metáfora associada a ele, estamos, naturalmente, identificando rupturas qualitativas para os produtos jornalísticos digitais. Tais rupturas são expressivas, pois modificam processos, criam procedimentos novos e propiciam novas funções ao instituir as bases de dados como definidoras da estrutura, da organização, bem como da apresentação dos conteúdos de natureza jornalística, de acordo com funcionalidades e categorias específicas para a criação, a manutenção, a atualização, a disponibilização e a circulação de produtos jornalísticos digitais dinâmicos. Esse paradigma, além de trazer possibilidades até então inexistentes, amplifica usos e características prévios, aperfeiçoa modelos e práticas, e garante inovação.

A generalização do uso de BDs em *sites* jornalísticos concorre para sobrepujar modelos transpositivos resultando, conseqüentemente, no distanciamento da metáfora do impresso. Marcos Palacios (2002a, 2003b, 2003c) conjecturou sobre a superação paulatina daquele padrão, à época apontando que a introdução de recursos de memória, a exemplo dos arquivos, associados às características de hipertextualidade, instantaneidade e interatividade representavam uma ruptura para o jornalismo digital em relação às modalidades anteriores.

Já Luciana Mielniczuk (2003) defendeu que o webjornalismo de terceira geração precisava apresentar rupturas suficientes para configurar os *sites* jornalísticos verdadeiramente como um novo produto, rompendo com padrões e oferecendo possibilidades inéditas. Consideramos que os *sites* jornalísticos estruturados em bases de dados efetivamente se distinguem e concretizam o ideal de apresentar rupturas qualitativas, agora numa fase de transição entre a terceira e uma quarta geração do jornalismo digital. No limiar, portanto, da segunda década de desenvolvimento desta modalidade. Deslindemos, então, algumas dessas rupturas e potencialidades engendradas pelo modelo **JDBD**.

### **5.1.1 Produto digital metainformativo, inteligente e dinâmico**

Esta é designação proposta para os *sites* jornalísticos que empregam as bases de dados. Denominamo-os como tal, tomando como parâmetro as funcionalidades e as categorias do **Jornalismo Digital em Base de Dados (JDBD)** e tendo como inspiração os estudos realizados pelo professor Raymond Colle (2000, 2002, 2003, 2005a, 2005b). O produto digital metainformativo, inteligente e dinâmico distingue-se radicalmente dos anteriores *sites* estáticos e têm na dinamicidade um dos pontos diferenciais que habilita a sua implementação, promovendo a adequação dos processos, com o uso de BDs, associadas a linguagens de programação, a recursos de metadados, *data mining*, entre outros.

Nos *sites* de tipo dinâmico, o conteúdo é permanentemente renovado, inter-relacionado, com apresentações diferenciadas para os mesmos, o sistema de busca é refinado, e os serviços complementares, como estatísticas dinâmicas, são ajustadas a

cada instante, à medida em que se ampliam os conteúdos nas bases de dados. Neles, o princípio básico de construção não é a página em si, mas consiste em unidades ou frações de conteúdo - textos, fotos, cliques de vídeo, áudio, infografias interativas, etc. -, que são alterados de acordo com um gabarito para criar uma página. O conteúdo é separado da forma graças à utilização da *XML*, dos metadados, das anotações semânticas sobre os dados para descrever o conteúdo, facilitando a identificação e a localização das unidades dos objetos digitais, dos gabaritos para descrever as características do conteúdo (tipo, tamanho, formato) e dos procedimentos para saber qual conteúdo muda dentro do gabarito. É o que Francesco Cara e Stefana Broadbent (2000) denominam arquitetura contextual, a qual oferece novas possibilidades para a criação dos *web sites* e fornecem aos usuários novos caminhos para descobrir e explorar os conteúdos (CARA; BROADBENT, 2000, p.206).

Uma vez estruturado em BDs, a gestão do produto como um todo passa a contar com a possibilidade de automatização das etapas do processo de apuração, produção, recuperação, composição, edição, apresentação e circulação dos conteúdos, permitindo aos profissionais dedicarem maior tempo ao trabalho intelectual e analítico, inclusive para pensar sobre critérios para indexação e classificação do material, assim como para o estabelecimento de relações entre as informações. Os sistemas de apuração, por sua vez, passam a ser menos hierarquizados e a produção descentralizada flexibiliza e facilita o trabalho dos profissionais.

Como a migração para uma estrutura em base de dados tem sido gradual e, em alguns casos, mais lenta, é possível às médias e pequenas empresas informativas, por exemplo, irem incorporando-as aos poucos, seja como fonte para a pesquisa e apuração, para regular o sistema de fontes, para disponibilizar os arquivos e para a integração deles às matérias e à oferta informativa, seja para apresentar os conteúdos.

Cabe lembrar que este produto incorpora, ainda, diferentes plataformas, como a web, a própria rede local das empresas ou intranet, *e-mail*, aplicações sem fio ou *Wireless Application Protocol*<sup>228</sup>, além de *chats*, fóruns, ferramentas para o gerenciamento do fluxo de trabalho, dentre outras. Também prevê a presença de jornalistas nas equipes de desenvolvimento dos sistemas de gestão de conteúdos que darão todo o suporte e conterão todas as especificações conceituais e técnicas para a produção, composição, edição e apresentação das informações.

### 5.1.2 Sistemas de produção, edição e circulação automatizados

O desenvolvimento de sistemas automatizados para o armazenamento, recuperação, distribuição e comercialização de informações jornalísticas foi inaugurado nos anos 70 do século XX justamente com o emprego das bases de dados. Os aperfeiçoamentos posteriores desta tecnologia e o seu potencial de reinvenção as fizeram atravessar décadas e chegar no século XXI como uma solução eficiente e amplamente adotada para a construção de *sites* dos produtos jornalísticos.

Como definidoras da estrutura, da organização, e da apresentação dos conteúdos, as BDs proporcionam a automatização para os processos jornalísticos, compreendendo desde as etapas iniciais de pré-produção, edição até a circulação e a repercussão que tais conteúdos possam alcançar na Web. Desta maneira, verifica-se uma ruptura com modelos e práticas anteriores, para a conformação de procedimentos e do produto em si, embora, por outro lado, co-existam a melhoria e o aprimoramento de atividades intrínsecas à prática, como a coleta, pesquisa, seleção, checagem, e análise das informações, recursos para a composição, a edição e os modos de representação visual

---

<sup>228</sup> *WAP* é um padrão internacional para aplicações que utilizam comunicações sem fio (internet móvel), como por exemplo acesso à web ou ao correio eletrônico a partir de celulares, entre outros dispositivos móveis.

para as peças informativas. Os metadados, o *data mining*, o *Knowledge Discover in Database (KDD)*, entre outras técnicas de prospecção e extração de conhecimento embutido no material armazenado em BDs trazem mudanças, favorecendo o aprofundamento, a contextualização e auxiliam na criação de tematizações diferenciadas e elementos conceituais novos para denominar seções, além do que instauram padrões menos convencionais para apresentar os conteúdos.

A automatização dos sistemas de produção, composição, edição e circulação ocorre segundo três tipos: parcial, quando ela contempla algumas etapas do processo de produção jornalístico; procedimental (nível intermediário), quando se tem mais etapas do processo jornalístico funcionando de modo automatizado; e total, quando os sistemas são totalmente automatizados, e ocorrem sem qualquer intervenção dos jornalistas.

Consideramos os agregadores de notícia – mesmo criticados por serem reempacotadores do conteúdo produzido originalmente por jornalistas de outros *sites* noticiosos - como os produtos nos quais as rupturas com relação a modelos anteriores estão mais evidenciadas - vide casos como os do *NewsBrief* e *NewsExplorer*, *Google News*, *Topix.net*, *Newsmap* e *Newzbubble* -, ainda que elas também se apresentem nos *sites* da *BBC News*, *Elmundo.es*, *Elpaís.com*, *The New York Times*, *Washingtonpost.com*, *Libération*, *MSNBC*, *Folha Online*, *International Herald Tribune*, *Portal Terra* e o blog *Deu no Jornal*.

Os que se enquadram na categoria de produtos acadêmicos laboratoriais igualmente romperam com padrões prévios (*TDC* e *Akademia*) e estão experimentando novas configurações, a exemplo do *Panopticon* e do *Urbi et Orbi*, este, inclusive incorporando o elemento vídeo na construção da própria narrativa e não de modo justaposto como é mais utilizado. Quanto à *Notibits*, revista digital com características de agregador, já que recompila notícias mais destacadas sobre novas tecnologias e

informática a partir da coleta das mesmas em *web sites* de outras fontes informativas, ainda se encontra num estágio mediano de automatização.

Pelo que se observa, a tendência aponta para o surgimento de mais agregadores, sejam eles totalmente automatizados ou sejam de tipo procedimental, integrando também jornalistas para realizar as tarefas de seleção, checagem e edição do material. Dois exemplos são: o *DailySource.org*<sup>229</sup>, organização sem fins lucrativos criada em 2005, operando ainda com versão beta e incorporando jornalistas no processo de seleção, e o *Daylife*<sup>230</sup>, um dos mais novos agregadores, lançado em janeiro de 2007, também disponibilizado em versão beta. Funciona de modo automatizado, prometendo oferecer uma maneira diferente de se explorar o mundo, segundo um ordenamento que privilegia a atualidade e a relevância dos assuntos, a partir de perspectivas variadas e com contextualização.

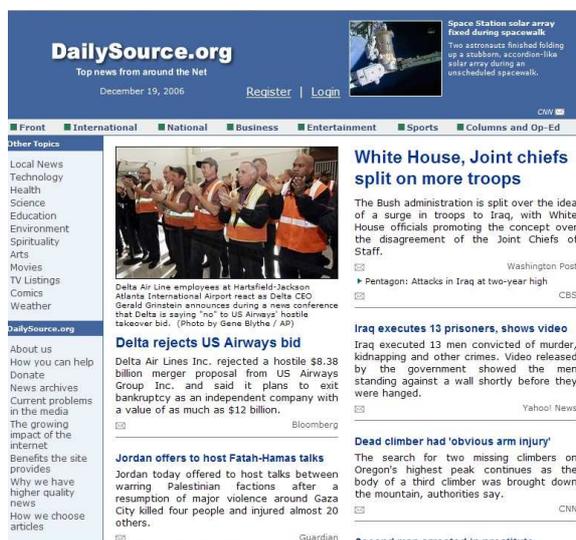


Figura 43 – *DailySource*: agregador incorpora jornalistas para o trabalho de seleção e edição e não possui anúncios publicitários. Jan.2007.

<sup>229</sup> <http://dailysource.org>.

<sup>230</sup> <http://www.daylife.com>.



Figura 44 – O *Daylife* promete maneira diferente para explorar o mundo, com perspectivas variadas e contextualização. Fev.2007.

Além dos agregadores, certamente produtos de outros perfis despontarão com as características que se aponta e com muitas outras que poderão trazer mais diferenciais.

### 5.1.3 Composição das peças informativas e as notícias de ordem superior

Conformar padrões novos para a construção das peças informativas corresponde ao emprego das bases de dados como estrutura e suporte para garantir a produção de narrativas interativas – a hipernarrativa - de maneira mais afinada com os recursos proporcionados pelo hipertexto, pelo espaço navegável, pelo suporte digital. Sendo assim, observamos a configuração de mudanças qualitativas para a narrativa jornalística e, baseada nas proposições de Lev Manovich (2001), Christiane Paul (2004), Elias Machado, (2004a, 2004b, 2006) e Beatriz Ribas (2005), inferimos que o uso das BDs em conjunto com as técnicas e linguagens associadas a elas e com a incorporação das propriedades do jornalismo digital aos sistemas de gestão de conteúdos, a partir dos *templates* criados, assegura outras possibilidades para, concretamente, gerar peças informativas mais originais.

Segundo António Fidalgo (2004, p.05), a feitura da notícia *on-line* mediante uma base de dados, apesar de responder às célebres perguntas O que, Quem, Quando, Onde, Como e Porque, não se adequa ao formato da pirâmide invertida<sup>231</sup>, devendo ser construída de um modo diferente. Diante disso, pensamos que o modelo da pirâmide invertida deitada, como sugerido por João Canavilhas (2005, 2006), permite organizar a produção das peças informativas em conformidade com a concepção de resolução semântica, que inclui as noções de alta e baixa resolução, e define as categorias de densidade informativa e diversidade temática para o **JDBD**.

A pirâmide invertida deitada contempla quatro níveis, partindo de um estágio de menos informação para o de mais informação, a ser seguido pelos usuários a partir de seu interesse: o nível 1 é a unidade base (*lead*, respondendo ao O que, Quando, Quem e Onde); o nível 2 é o de explicação (responde ao Por que e ao Como); o nível 3 é de contextualização (oferecendo mais informação, seja textual, vídeos, áudios, infografias interativas, *slideshows*, etc.); e o nível 4, de exploração (que liga a notícia ao arquivo da publicação ou a arquivos externos). Os quatro níveis podem, inclusive, auxiliar numa divisão da produção de conteúdos em redações digitais, e ser utilizados para a definição das permissões no acesso aos sistemas de gestão de conteúdos. Evidentemente que para as operações onde a convergência já está em curso, a aplicação dos níveis se dará de maneira integrada.

---

<sup>231</sup> Tal arquitetura noticiosa, que estabelece que uma notícia começa pelos dados mais importantes seguido de informações complementares organizadas em blocos decrescentes de interesse, nasceu durante a Guerra da Secessão, nos Estados Unidos, em 1840, com o telégrafo, a grande inovação tecnológica da época (FONTCUBERTA, 1999, p.58). Foi batizada como pirâmide invertida por Edwin L. Shuman no livro *Practical Journalism* (SALAVERRÍA, 2005a, p.109), tornando-se uma das regras mais conhecidas no meio jornalístico. A validade de tal técnica vem sendo alvo de questionamentos e críticas por transformar o trabalho jornalístico numa rotina, deixando pouco espaço para a criatividade. Por outro lado, desde o surgimento do jornalismo digital, tem motivado polêmicas, dividindo os que ainda a consideram como um modelo válido, e aqueles que defendem que ele não serve à esta modalidade. Ainda que se perceba a sua utilidade para a redação das *breaking news*, por exemplo, a pirâmide invertida se revela inadequada para compor outros gêneros. Sobre o assunto ver: Díaz Noci, 2001; Díaz Noci; Salaverría, 2003; Mielniczuk, 2003; Salaverría, 2005b; Edo, 2005; Canavilhas, 2005, 2006.

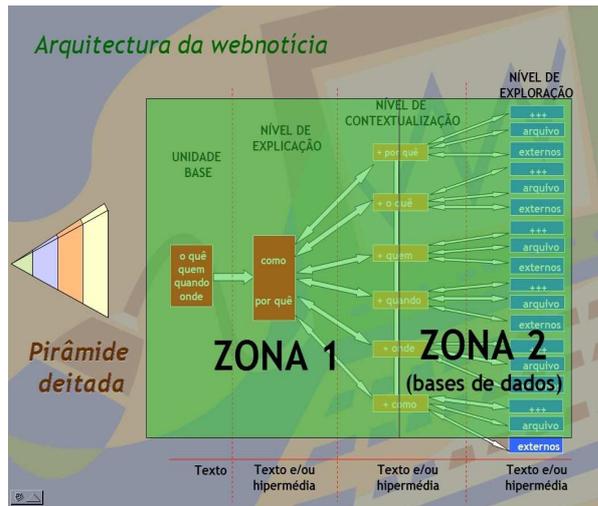


Figura 45 – Técnica da pirâmide deitada contempla as BDs nos níveis de contextualização e exploração, linkando para os arquivos da publicação e para *sites* externos.

Fonte: Canavilhas (2005, 2006).

Cabe notar que, ao indicar o nível de contextualização e de exploração, Canavilhas prevê nestas etapas o uso de material armazenado, e organizado nas bases de dados para oferecer maior complementação das informações, com documentação e dados extraídos a partir da co-relação entre as diversas notícias e reportagens, e também com a integração de arquivos externos ou com a inserção de *links* de remetimento para outros *sites*. Portanto, diferentemente da pirâmide invertida “clássica”, a pirâmide invertida deitada não implica em interesse decrescente no texto noticioso, mas, pelo contrário, em níveis de aprofundamento e contextualização à medida que se navega em direção à base da pirâmide.

A extração de informações pode contemplar, ainda, os metadados, o *data mining* e o *KDD* para descobrir dados novos, bem como fatos, relações, comportamentos, tendências, perfis, de grande valor noticioso, a partir da prospecção, pesquisa e análise. Neste sentido, o produto metainformativo, inteligente e dinâmico estará trabalhando com notícias de ordem superior, conforme conceitualização proposta por António Fidalgo (2005). Referindo às notícias de observação direta como notícias de primeira

ordem ou de ordem normal, ele esclarece que as notícias de ordem superior são aquelas notícias obtidas através da mineração de dados, resultando num novo jornalismo investigativo. Para a formulação do conceito designativo deste tipo de notícias, Fidalgo aplica, em paralelo, a teoria dos objetos de ordem superior de Alexius Meinong<sup>232</sup>.

A estruturação da notícia segundo os parâmetros indicados assegura uma narrativa criativa, contextualizada, interativa e multimídia, agrupando recursos hipertextuais, de áudio, de vídeo, as infografias interativas, até animações tridimensionais, o que concorre para tornar a experiência dos usuários com os conteúdos jornalísticos mais envolvente, autêntica, interessante e imersiva (PAVLIK, 2005, p.48-49). Por outro lado, agrega o sentido de agência, um dos prazeres característicos dos ambientes eletrônicos, como afirma Janet H. Murray (1997, 1999, 2003a): “Agência é a capacidade gratificante de realizar ações significativas e ver os resultados de nossas decisões e escolhas. Quanto mais bem resolvido o ambiente de imersão, mais ativos desejamos ser dentro dele” (MURRAY, 2003a, p.127). Cabe lembrar ainda que, conforme explicado por Elias Machado (2004b, p.05), o fluxo da narrativa no ciberespaço depende da intervenção direta do usuário para manter a sua qualidade, ao mesmo tempo em que ativa as relações estabelecidas por ele com o conteúdo, com a publicação e com os demais atores.

Para além dos recursos tecnológicos que promovem a imersão, a incorporação das poéticas da imersão tal como apresentadas por Marie-Laure Ryan (2001, p.120-171)

---

<sup>232</sup> A ciência da objectologia do filósofo alemão Meinong foi tema da tese de doutorado de António Fidalgo, a qual resultou no livro *O realismo da fenomenologia de Munique* (1991). “A objectologia é a ciência universal por excelência na medida em que tudo é objecto. Representações, percepções, juízos visam sempre algo e este algo é um objecto. Neste sentido, as montanhas de ouro e os círculos quadrados são tão objectos como os objectos reais com que lidamos no dia a dia”, explica Fidalgo (1991, p.244). Os objetos, segundo Alex Meinong, são classificados em quatro grandes classes principais, definidas por ele a partir das vivências psíquicas elementares de representar, pensar, sentir e desejar, correspondendo, então, às classes dos “objecta”, “objectiva”, “dignitativa” e “desiderativa”. Os objetos de ordem superior pertencem à classe dos “objecta”, assim como os objetos de ordem inferior. “Os primeiros fundam-se nos segundos. Assim, a diferença é um objecto de ordem superior fundado nos objectos que são diferentes (...) Relações e complexões são em geral objetos de ordem superior” (Ibidem:., p.249).

também podem colaborar. Baseada nas teorias literária e de realidade virtual, a autora define imersão como uma experiência corporal que toma a projeção do corpo virtual ou do próprio corpo para sentir-se integrante de um mundo artístico. Para Ryan, imersão e interatividade são as duas dimensões da realidade virtual. A interatividade, por sua vez, envolveria as habilidades do usuário para modificar o ambiente durante uma simulação, com movimentos do seu corpo.

Poder-se-ia, então, considerar que um usuário ao interagir com conteúdos jornalísticos, participando como colaborador no processo de produção das informações, estaria desempenhando uma função de co-autor, o que também está conectado com a idéia de um jornalismo participativo. Deste modo, ao nos apropriarmos das poéticas da imersão descritas por Ryan (2001), teríamos:

- Imersão espacial. O usuário desenvolve um senso de lugar, um senso de fazer parte da cena dos eventos narrados. Ou seja, conteúdos mais contextualizados, construídos com uso do hipertexto, áudio, vídeo, fotos, infografia interativa; participação dos usuários na alimentação do conteúdo, contribuindo para a diversidade temática;

- Imersão temporal. A experiência do usuário é envolvida pelo suspense da narrativa, o desejo de saber o que acontece depois. Quanto mais o usuário for motivado a “entrar” numa notícia ou reportagem pelo que a sua narrativa possa conter de atrativo, mais se sentirá motivado e interessado a acessar determinado produto;

- Imersão emocional. O fenômeno de desenvolver uma conexão pessoal com os personagens. Poder interagir com os jornalistas e o produto, opinando, sugerindo e publicando seus próprios conteúdos, certamente, proporcionará maior interatividade entre os usuários e a publicação. Conseqüentemente, laços emocionais e uma relação de maior proximidade poderão ser estabelecidos.

A adoção de BDs, portanto, favorece a inovação para a construção das peças informativas, segundo parâmetros específicos, permite incorporar técnicas para ampliar a descoberta de fatos novos, explorar inter-relacionamentos entre as informações, apresentar os conteúdos de maneira menos convencional, assegurando uma oferta informativa mais diversa.

#### **5.1.4 Arquivos como “bases de conhecimento”**

O armazenamento das informações jornalísticas em bases de dados computadorizadas - processo iniciado na década de 70 e que tem como um dos marcos o *The New York Times Information Bank* - mudou a maneira como as organizações jornalísticas até então tratavam os “*morgues*”, as BDs analógicas que, embora referenciais para as empresas e para o público externo, necessitavam de melhor organização e indexação para a recuperação mais ágil das informações. A partir dos anos 70, as organizações jornalísticas passaram a valorizar mais os arquivos, transformando-os em peças de alto valor agregado, seja para a pesquisa, para a recuperação e contextualização das informações, seja para impulsionar as receitas, com a distribuição e a comercialização desses conteúdos.

A implantação dos departamentos de documentação, com os arquivos eletrônicos, favoreceu a migração para uma plataforma digital e facilitou a trajetória de muitas empresas informativas na web, colocando, por outro lado, exigências cada vez maiores quanto à precisão na redação e nas análises dos acontecimentos. Tal precisão resulta no melhor aproveitamento do material, na sua recuperação para a pesquisa, na interpretação jornalística dos fatos, na reconstrução histórica e no desenvolvimento da visão de conjunto.

Com o avanço na digitalização dos arquivos jornalísticos e a sua disponibilização nos respectivos *sites* dos produtos, cria-se para os usuários mais condições de interação com as organizações noticiosas. Essas, por sua vez, promovem o resgate do passado e o colocam a serviço do público, cumprindo a sua função para a atualização da memória social de uma maneira instantânea, múltipla e hipertextual.

Os arquivos *on-line*, constituídos, de modo geral, pelas edições publicadas por um produto jornalístico (FIDALGO, 2003) e organizados em bases de dados, geram conhecimento, sendo que seu aprimoramento e aplicação para a pesquisa, análise, interpretação e contextualização lhes asseguram a designação de “bases de conhecimento” (COLLE, 2000, 2002).

Examinando-se a maneira como eles têm sido incorporados à oferta informativa, na estrutura da própria notícia, da reportagem, ou para gerar novos eixos temáticos para a organização conceitual dos conteúdos (“O que fez esse dia entrar para a história”, “O que já foi notícia”, “On this day”, “A fondo”, etc) constata-se uma outra inovação para os produtos digitais, relacionada com a característica da memória, que por si só já configura uma ruptura para o jornalismo digital frente às modalidades tradicionais (PALACIOS, 2002a, 2003b, 2003c, 2004).

Usos editoriais distintos do material de arquivo (textos, fotos, vídeos, áudios infografias interativas, etc.) podem ampliar ainda mais a diversidade temática, por meio da co-relação entre as informações garantida pela flexibilidade combinatória. Por outro lado, os inter-relacionamentos entre conteúdos publicados anteriormente por um mesmo *site* ou os que remetem para os de *sites* externos acrescentam maior profundidade e oferecem perspectivas diferenciadas quanto ao tratamento dos eventos noticiosos.

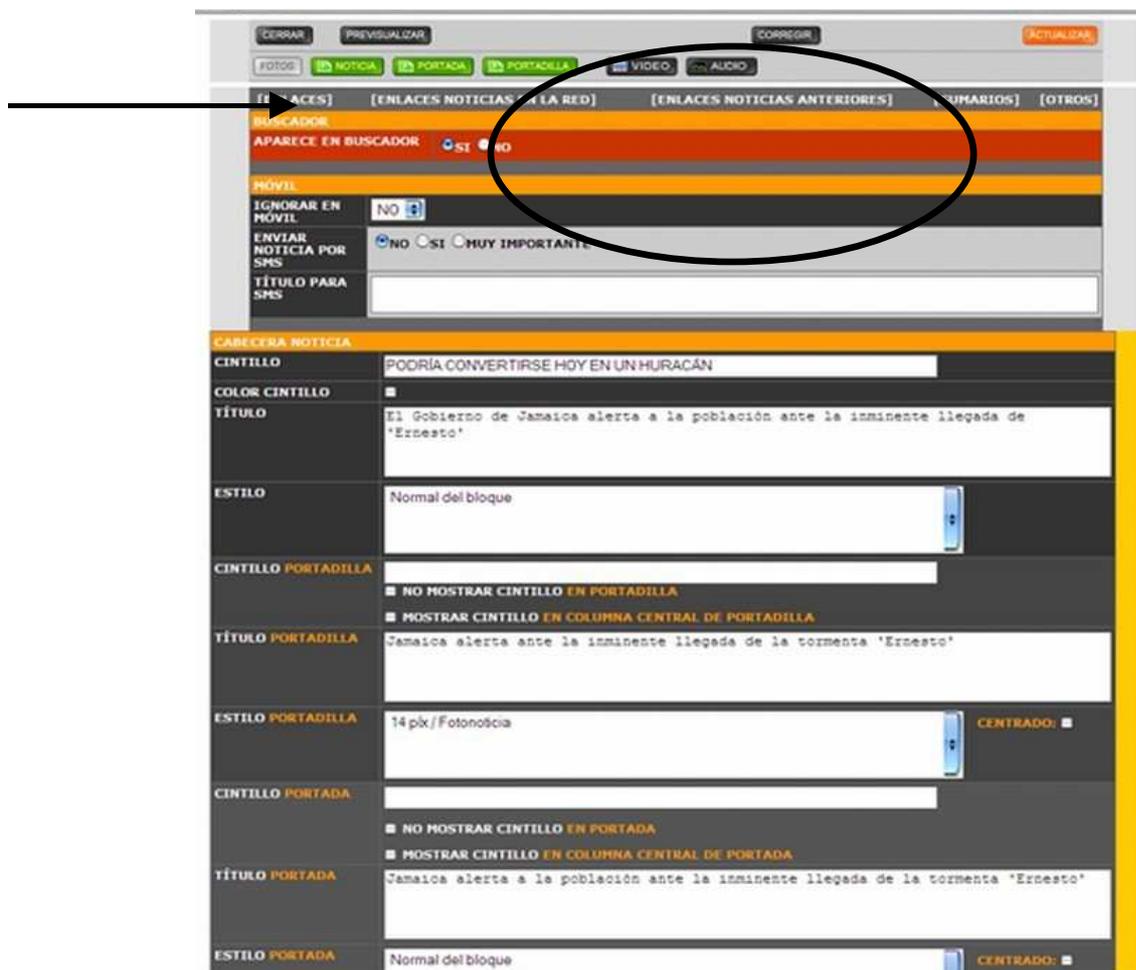


Figura 46 – Gestor de conteúdos do *Elmundo.es*. Marcadores indicam os botões de inserção para *links* de inter-relacionamento apontando para o arquivo do produto, bem como para o de outros *sites*.  
Fonte: Rivero, 2006.

Ao explorar novas tematizações e elementos conceituais para a classificação externa do material de arquivo, surge, em contrapartida, o potencial para que ele conforme um novo gênero ou subgênero. Tal hipótese é colocada tendo como base as fases de evolução para os gêneros no jornalismo digital, conforme trabalhado por Ramón Salaverría e Rafael Cores (2005), nomeadamente os níveis de: repetição (nível mais básico, o da reprodução dos gêneros já existentes nos meios tradicionais), o de enriquecimento (incorpora possibilidades hipertextuais, multimídias e interativas), o de renovação (recria gêneros precedentes, reconfigurando, por exemplo, a notícia e a reportagem) e o de inovação (cria gêneros para os produtos digitais como um modelo de anúncio específico, socialmente reconhecível). Os dois últimos são os níveis que nos

permitem indicar os arquivos como um potencial gênero ou subgênero e, em paralelo, apontar o enriquecimento e a renovação que eles possuem para recriar outros gêneros.

### 5.1.5 Evolução nos modos de apresentação das informações

Tomando como referência os trabalhos de Mindy McAdams (1995), C.R. Watters, M.A. Sheperd, T. Chiasson, L. Manchester (1997, 2000), Lars Bo Eriksen & Carina Ihlström (2000), Maria Åkesson (2003), Carina Ihlström e Maria Åkesson (2004) e o de Jonas Lundberg e Carina Ihlström (2004), que analisaram os procedimentos de estruturação, bem como a apresentação dos conteúdos das edições digitais de jornais norte-americanos e suecos - amparados em pesquisas sobre o comportamento do usuário ao navegar por esses *sites* -, para confirmar a *broadsheet metaphor* como modelo apropriado, e o estudo desenvolvido nesta tese, consideramos o **Jornalismo Digital em Base de Dados (JDBD)** com o potencial para fazer emergir outros padrões para guiar a apresentação das informações jornalísticas. Esses padrões se distanciam daquele baseado no impresso e dão forma a uma nova metáfora, ancorada na estética base de dados ou *database aesthetics*.

Como tal, ela expande as opções de uso de técnicas para a visualização das informações e conformação do *layout* dos produtos, traduzindo em diversidade e originalidade para a criação de interfaces<sup>233</sup>. Ademais, delimita uma estratégia comunicacional (GROSSMANN, 2006) mais apropriada à produção jornalística no

---

<sup>233</sup> Segundo Xosé Pereira (2006, p.214), a construção de interfaces, seja de computador, ou seja da web, deve sempre guardar taxas suficientes de metáfora que permitam uma fácil assimilação por parte dos usuários. O emprego de metáforas e a sua importância para a criação de interfaces tem sido abordado por diversos autores, entre eles Carlos Scolari (2004), o qual aporta relevante categorização para as metáforas digitais, o próprio Pereira (Ibid.), que destaca a presença gráfica do sistema *front end* como fundamental para alcançar o êxito comunicativo, além dos estudos referenciados neste capítulo sobre a metáfora do impresso, entre outras, e a sua aplicação nos produtos jornalísticos digitais.

ciberespaço. A proposição desta metáfora se conecta, por outro lado, com as avaliações e sugestões indicadas por aqueles investigadores, que, em última instância, identificaram a necessidade de evolução para os padrões de apresentação das informações, colocando a flexibilidade como uma vantagem útil para adequar o formato de representação dos conteúdos com as ações requeridas dos usuários para explorá-los.

Como afirmado na abertura deste capítulo, esta nova metáfora para produtos digitais, metainformativos, inteligentes e dinâmicos não determinará um padrão único, visto que técnicas e métodos diversos podem ser utilizados para gerar visualizações diferenciadas, conforme é possível inferir a partir dos exemplos empíricos trabalhados ao longo da tese.

Em geral, os casos analisados possuem, especialmente nas *front pages*, aspectos visuais descolados do estilo colunado do impresso, por exemplo, com os destaques para os assuntos principais também arranjados em modos diferenciados. No entanto, a divisão das seções e os elementos conceituais utilizados para organizar a classificação externa das informações ainda reproduzem aquelas do impresso, tanto nos produtos vinculados a operações convencionais, como naqueles essencialmente digitais e até mesmo nos que são totalmente automatizados. Isso demonstra a força da metáfora do impresso e do eixo temático tradicional - política, economia, esportes, cultura (FONTCUBERTA, 2006), além de ciência, saúde, tecnologia -, para compartimentalizar os conteúdos, embora outras tematizações sejam, simultaneamente, trabalhadas.

O *site* do *International Herald Tribune*, por exemplo, é um dos que mantêm mais fortemente a metáfora do impresso, tanto na *home* como na estrutura das peças informativas, ainda que associada a outros padrões visuais e a recursos de dinamicidade que criam maior interatividade dos usuários com o produto. Nos demais, o formato das

peças informativas segue o padrão da metáfora do documento web (WATTERS et al, 1997; 2000, p.04) - isto é, apresentando o elemento textual em notícias e reportagens no modo típico de documento único WWW, ao lado dos recursos de áudio, vídeo, *links de relacionamento*, etc., – e requerendo a necessidade de rolagem da barra lateral.

Pacotes informativos essencialmente multimídias, infografias interativas, listas ordenadas, seções de estatísticas dinâmicas, como “Most popular”, “+ lidas”, tabelas de páginas (*page rank*) e aquelas opções que permitem explorar relações que se estabelecem em torno de personagens e atores das notícias, são algumas das peças nas quais novos métodos de visualização estão sendo aplicados. Já no *Newsmap* e no *Newzbubble*, os métodos *treemap* e o *data mining* são empregados para gerar as interfaces, nas quais se identificam as visualizações mais inovadoras para *front pages*.

Como a migração para uma estrutura em BD é crescente, aumenta, paralelamente, o alcance de aplicação e expansão desta metáfora **JDBD**, fundamentada na estética base de dados, para os produtos digitais. Neste sentido, pode conduzir a um cenário de potencialização em grau elevado desta metáfora que deverá resultar no desprendimento total com o padrão da *broadsheet metaphor*, abrindo o caminho para a criação de outras metáforas efetivas.

### 5.1.6 Elementos conceituais para a organização do conteúdo

A análise do *corpus* empírico nos permitiu encontrar elementos conceituais novos para a classificação externa e a conseqüente organização dos conteúdos que compõem a oferta informativa dos *sites*, sendo possível constatar as inovações trazidas pelo paradigma **JDBD** também quanto ao eixo temático. Desta maneira, localiza-se uma

outra mudança qualitativa em relação aos parâmetros que anteriormente definiam a organização dos conteúdos nos produtos jornalísticos digitais.

Tais inovações estão atreladas ao uso de algoritmos, assim como de tecnologias específicas desenvolvidas pelas equipes de alguns produtos, da incorporação de recursos como *RSS*, *podcasting*, *vídeo*, *áudio*, metadados, *data mining*, *tags*, *links* de inter-relacionamento, que ampliam as possibilidades para se oferecer conteúdos mais originais e criativos, favorecendo a densidade informativa e a diversidade temática. As opções para compartilhamento e repercussão dos conteúdos na WWW, principalmente nos *blogs* e nos chamados *sites* sociais, são outros diferenciais.

Como ilustrado pelo quadro a seguir, tem-se novos elementos classificando o material de arquivo, o que é elaborado pelos usuários, o que é gerado de modo automatizado - como as listas ordenadas e estatísticas dinâmicas -, além daquele que provém das co-relações entre eventos, pessoas, entidades, organizações. Há, ainda, os conteúdos que exploram temáticas ficcionais e de entretenimento, como novelas, *games* e *cartoons*.

As potencialidades para as novas tematizações, bem como para os elementos conceituais são certamente maiores, uma vez que a tendência atual privilegia a informação cada vez mais estruturada, mais pautada e organizada para permitir o estabelecimento de relações e inter-relacionamentos. Portanto, algo que é intrínseco ao emprego de bases dados para construir os *sites* dinâmicos.

<b><u>ELEMENTOS CONCEITUAIS</u></b>		<b><u>TECNOLOGIAS E RECURSOS AGREGADOS</u></b>
PARA MATERIAL DE ARQUIVO	On this day, A fondo, O que já foi notícia, O que fez esse dia entrar para a história, TimesSelected, Arquivo	<b>Algoritmos, tecnologias proprietárias, RSS, Metadados, Podcasts, algoritmos, data mining, KDD, Vídeos, Áudios, Tags</b>
PARA PARTICIPAÇÃO DO USUÁRIO	Jornal do cidadão, vc repórter, vc diretor, Have your say, First Person, Your Stories, Participa, Vós réactions, Vos photos sur Libé, La question du jour, Post a comment, Recent Comments, Op-Ed Contributors	<b>Links de Inter-relacionamento incorporados às peças informativas:</b> “Dans la même rubrique”, “Sur le même sujet”, “Noticias relacionadas en elmundo.es”, “ Noticias relacionadas en otros medios”, “La noticia en otros webs”, “Webs en español”, “En otros idiomas”, “More on Washingtonpost.com”, “Related Articles”, “Recommended Stories”
ESTATÍSTICAS DINÂMICAS E LISTAS ORDENADAS	Most Popular, Les +, + Lidas, + Comentadas, Most Viewed Articles, Lo más visto, Lo último, Lo más leído, ... enviado, el indice, Most Emailed Articles, Top MSNBC, Top Rated, Recent Comments	<b>Opções de Compartilhamento e Repercussão dos conteúdos:</b> “Who’s blogging? – Read what bloggers are saying”, “Most Blogged About Articles”: On Washingtonpost.com; On the web (Technorati)”, “Share”, “Digg”, “Del.icio.us”, “Google”, “Yahoo”, “Redict +”, “Facebook”, “My MSN”, “Bloglines”
OUTROS	Related People, Related Organizations, Countries, Timeline, Blogs, Offbeat, News Alerts, Mobile News, Recommended, AudioNews, Videos, Pacotes multimedia: Faces of The Fallen, Faces of The Dead , Portraits, Grands angles, Tentations, Votes Databases, Gráficos (para infografias interativas), Special Report, Mi País, Local, Nation, <i>Equilibrio, Futuro</i>	
CONTEÚDO FICCIONAL E DE ENTRETENIMENTO	Novelas - como Jezebel’s Tomb - Games, Cartoons	

Quadro 7 – Elementos conceituais que organizam a classificação externa dos conteúdos, tecnologias e recursos agregados à oferta informativa

### 5.1.7 O jornalista, as competências cognitivas, as novas funções

Atualmente, os jornalistas, entre outros profissionais da comunicação, trabalham imersos em ambientes caracterizados pelo forte cunho tecnológico, influenciando todas as etapas do processo produtivo, independente do produto final que seja entregue ao público. Esses ambientes são ainda mais complexos se são destinados à produção de *web sites* jornalísticos, sendo, então, correto afirmar que o aspecto tecnológico e da presença do digital são altamente potencializados nas redações de produtos digitais.

Diante desse cenário é que se considera possível classificar tais espaços como redações inteligentes, nas quais o produto metainformativo inteligente e dinâmico é gerado. Tal denominação poderia ser estendida, ainda, às redações onde a convergência já fundiu os espaços de trabalho de um diário impresso e a sua edição *on-line*, por exemplo, pois também dependem fortemente de implementações tecnológicas para funcionarem de modo integrado.

Para atuar nessas redações, é cada vez mais requerido do jornalista habilidades cognitivas para desempenhar as funções, operando com sistemas de gestão de conteúdos automatizados e, inclusive, atuando com protagonismo nas equipes de desenvolvimento dos mesmos. Para além da formação acadêmica, esse profissional deve aprender a aprender e perseguir a qualificação continuada, já que nem sempre as organizações jornalísticas investem recursos para a sua formação.

Assim, tendo como referência as competências para o comunicador digital conforme definidas pelo estudo produzido pela Rede Ibero-Americana de Comunicação

Digital (Rede ICOD), Projeto “Comunicadores Digitais”<sup>234</sup>, ao jornalista é recomendado: manter-se atualizado no uso de *softwares*, conhecer o potencial da tecnologia de bases de dados e a sua aplicação no jornalismo, as linguagens de programação, os recursos para difundir e compartilhar conteúdos, as técnicas para investigar novas formas de produção e prospecção de informações como metadados e a mineração de dados, compreender o entorno que o circunda e as questões relacionadas com a cibercultura, entender o funcionamento e as potencialidades das redes sociais e comunidades virtuais, assimilar as características do jornalismo digital para criar conteúdos, conhecer as bases do jornalismo participativo, as diferentes plataformas tecnológicas (web, intranet, *e-mail*, aplicações sem fio, etc.) e ambientes como *wiki*, os *blogs*, saber valorizar as TICs criticamente, compreender a legislação vigente, bem como as situações de regulação jurídica relativas aos direitos de autor, delitos, privacidade, *copyleft*, *creative commons*<sup>235</sup>, *software* livre.

Esse conjunto de saberes e práticas precisa ser, evidentemente, assimilado pelas organizações jornalísticas, pois, do contrário, podem ficar fadadas a uma subutilização dos recursos. E, deste modo, inviabilizar a efetiva incorporação de tecnologias, processos e equipes especializadas para a melhor implementação de seus produtos digitais.

---

<sup>234</sup> A Rede ICOD surgiu através da iniciativa de um grupo de investigadores da América Latina e da Europa (Universitat de Vic, Espanha, Universidad Nacional de Rosário, Argentina, Universidade da Beira Interior, Portugal, Universidad de la Patagônia Austral, Argentina, Universidad de La Habana, Cuba, Universidade Federal da Bahia e Universidade Regional do Noroeste do Estado do Rio Grande do Sul, Brasil, Université de Lille 3, França e a empresa italiana *Ars Media*). O projeto foi financiado pelo Programa Alfa da União Europeia, e vigorou de 2003 a 2005, com os resultados sendo apresentados em 2006. Rede ICOD: <http://www.icod.ubi.pt>.

<sup>235</sup> <http://creativecommons.org>. Um sistema alternativo às licenças tradicionais de utilização de obras protegidas, que permite ao criador de uma obra decidir quais os direitos que pretende reservar para si, enquanto autoriza o público a trabalhar com base nas suas idéias. No modelo “Autoria Comum”, do *Creative Commons*, trabalha-se com “Alguns Direitos Reservados” ou *Some Rights Reserved*, em oposição ao modelo “Reservados Todos os Direitos” ou *All Rights Reserved*, que rege a existência de *copyright*.

Na base de tudo, é importante destacar, estão os códigos deontológicos e as regras que regem o jornalismo, uma atividade profissional regulada por legislação específica, que consiste no exercício de funções de pesquisa, recolha, seleção e tratamento de fatos, notícias ou opiniões, através de texto, imagem ou som, destinados à divulgação informativa pela imprensa, por agência noticiosa, pela rádio, pela televisão ou por outra forma de difusão eletrônica (GRADIM, 2005) e digital.

Ao lado das competências, somam-se outras exigências, que, segundo Maria Angeles Cabrera (2005, p.320), incluem: maior responsabilidade para trabalhar nos ambientes das redes descentralizadas, da imediatez informativa, que faz aumentar ainda mais os cuidados com a comprovação, verificação, e complementação dos dados e fontes, com a pluralidade dos conteúdos, com ampliações de vozes emissoras (por meio de *blogs*, “*sites* sociais” e outras formas alternativas), com a abertura da agenda informativa para atender a novas demandas dos usuários, especialização informativa para se ter mais rigor sobre o que informa, melhor uso da capacidade de moderação e de seleção e, ainda, ter clara a decisão sobre se quer ser empregado dos grandes meios, de produtos alternativos, ou ter seu próprio *site* ou *blog*.

Além das conhecidas funções atribuídas aos jornalistas, surgem outras novas, como as apontadas por Cabrera (2005), Elias Machado (2006) e o relatório da Rede ICOD. Quais sejam: gestor de conteúdos ou gestor da informação, que compreende as atividades de valorar, selecionar, classificar, organizar, estruturar, apresentar os conteúdos, adaptando-os dentro de cada meio no lugar mais adequado; administrador de notícias, que gerencia o processo de produção de informações para múltiplas e variadas plataformas nas redações altamente digitalizadas; narrador multimídia, que tem por função elaborar matérias multimídia através da combinação de diferentes elementos informativos, como textos, fotos, gráficos, vídeos ou animações; pesquisador de

notícias, que fornece as informações complementares para os narradores multimídia, disponibilizando textos, fotos, gráficos, vídeos ou animações para a composição dos produtos distribuídos para diferentes plataformas de distribuição; *web editor*, aquele que é responsável pela condução do *site* ou edição digital.

Além dessas funções, os jornalistas também podem ter protagonismo nas equipes de desenvolvimento dos agregadores totalmente automatizados, definindo critérios conceituais para os produtos, para a seleção de fontes, para a classificação e indexação das informações (tanto na inserção delas nas BDs como na classificação externa), para as relações que podem ser estabelecidas entre elas, para as tematizações, entre outros.

Para fechar este subtópico, elenca-se algumas recomendações para os jornalistas que trabalham ou ocuparão posições nas redações inteligentes para dar forma aos produtos digitais dinâmicos. As competências cognitivas devem habilitá-los a utilizar os sistemas de gestão de conteúdos automatizados para: a apuração e a recuperação da informação, assim como para a criação, composição e edição de conteúdos mais originais, utilizando o espaço navegável como suporte e ativando ambientes diferenciados para constituir experiências distintas para a exploração dos produtos jornalísticos (MANOVICH, 2001; MACHADO, 2004a, 2004b; 2006; PAUL, 2004; RIBAS, 2005); a construção de peças informativas, segundo o modelo da pirâmide invertida deitada (CANAVILHAS, 2005, 2006), o qual comporta uma variedade de possibilidades, incluindo os recursos hipertextuais, multimídias, os atributos dos ambientes eletrônicos e as poéticas da imersão (MURRAY, 1997; 1999; 2003; RYAN, 2001), para a elaboração do material noticioso; a utilização de metadados, da mineração de dados ou *data mining*, e do *KDD*, que conformam um novo jornalismo investigativo,

com as notícias de ordem superior; estimular, envolver e efetivar a participação dos usuários no processo de produção dos conteúdos jornalísticos.

### **5.1.8 Usuários atuando como fonte informativa**

A integração dos usuários no processo de produção dos conteúdos no jornalismo digital é mais um aspecto que o modelo **JDBD** ajuda a impulsionar. Uma vez trabalhando com uma estrutura em base de dados, é possível ordenar e qualificar os colaboradores e os chamados “repórteres cidadãos” de maneira mais efetiva.

Incorporar os usuários como fonte informativa significa partilhar com eles um pouco do controle que sempre esteve nas mãos das empresas informativas e, por conseguinte, dos jornalistas. Isto é, estabelecer um relacionamento mais próximo entre proprietários e usuários, desenvolvendo laços de pertencimento, despertando o interesse para que eles atuem com maior protagonismo e contribuindo, assim, para a criação de produtos mais atraentes e afinados com as suas expectativas.

Embora este seja um aspecto apontado como fundamental do jornalismo nas redes desde os anos iniciais dessa modalidade, foi tratado sem muita atenção por boa parte das organizações jornalísticas, principalmente as convencionais. Essa realidade se alterou mais concretamente nos últimos quatro anos com a expansão das comunidades *on-line*, dos *blogs*, dos *sites* e das redes sociais, fator que obrigou as empresas a readequarem seus posicionamentos e a definirem estratégias, de acordo com as respectivas linhas editoriais, para incorporar o material produzido pelos usuários na sua

oferta informativa<sup>236</sup>. Por outro lado, colocou para os jornalistas uma nova realidade e o desafio de aprender a dividir o poder com os usuários, compreendendo que eles também passariam a moldar as suas peças informativas.

Para Dan Gillmor (2004, 2005), esse é o aspecto que transforma o jornalismo em um meio mais profundamente cívico e democrático do que ele foi durante todo o século XX.

Aprenderemos que fazemos parte de um mundo novo e que os nossos leitores/ouvintes/telespectadores começam a ser parte do processo. Por exemplo: aceito que os meus leitores saibam mais do que eu – e este é um facto libertador, que não ameaça a minha vida de jornalista. Todos os repórteres, de qualquer ramo, deviam aceitar essa ideia. Usaremos as ferramentas do jornalismo cívico ou passaremos à história. O nosso código de valores, que inclui a necessidade de ser verdadeiro e imparcial, nunca deixará de ser importante e continuaremos, de certa maneira, a ser os guardiões do templo, mas a nossa capacidade de dar forma a grandes debates de ideias – e de as analisar – tornar-se-á pelo menos tão importante como a nossa capacidade para recolher os factos e para os relatar (GILLMOR, 2005, p.16).

A implementação de espaços e seções para a publicação do material produzido pelos usuários (textos, fotos, vídeos, além de comentários, sugestões e críticas) nos produtos jornalísticos digitais tem ocorrido de maneira mais efectiva, assim como têm sido incorporadas opções de compartilhamento dos conteúdos dos *sites* noticiosos em *blogs*, no *Digg*, e em outros meios sociais, que ativam ainda mais a interatividade e novos modos para se explorar o produto. Em adição, tem-se também as opções de

---

<sup>236</sup> Vale lembrar que tais estratégias não são exclusivas dos produtos digitais. Os jornais, os telejornais, revistas, além de outros produtos veiculados em distintos meios, vêm abrindo canais para que os usuários sugiram pautas sobre assuntos que gostariam de ver abordados, por exemplo, existindo também aqueles cujo conteúdo é produzido totalmente pelo público. A revista *Sou+eu* (<http://semanais.abril.com.br/soumaiseu/index.shtml>), lançada em novembro de 2006 pela editora Abril, por exemplo, tem seu conteúdo produzido pelas leitoras femininas, o público-alvo da publicação. Elas podem receber de R\$50,00 a R\$500,00 pelos relatos, R\$100,00 por receitas que, após testadas pela equipe de redação, sejam publicadas, ou R\$200,00, caso sejam escolhidas para estampar a capa da revista. Todo o material enviado pelas leitoras é checado pelos jornalistas que integram a equipe de redação, obedecendo aos critérios jornalísticos de: interesse, relevância, veracidade e fidedignidade das fontes. Um código de ética também foi criado para regular as relações entre usuários e publicação.

personalização, as quais permitem acessar ou descarregar os conteúdos que se deseja, assim como anúncios publicitários e classificados.

Nos casos analisados, detectamos que a maior parte deles trabalha com espaços para integrar os usuários, embora, por outro lado, se verifique que em produtos como o inovador *Elmundo.es*, e a *Folha Online*, ainda não haja seções específicas para abrigar a produção deles. No *Portal Terra*, existem duas seções destinadas ao material produzido pelos usuários: “vc repórter”, que abriga textos e fotos, e “vc diretor”, um concurso para escolha de vídeos, os quais passam pela avaliação prévia de funcionários do portal para verificar se estão de acordo com o regulamento e, depois, são disponibilizados para que os usuários do *site* elejam os melhores. Enquanto o *Elpaís.com* possui a seção “Participa”, entre outras opções de interatividade. No *The New York Times*, são aceitos artigos e *cartoons* animados, por exemplo, como contribuições em “Opinion” na subseção “Op-Ed Contributors”, além de cartas e comentários dos leitores, existindo ainda a figura do editor do público, que na subseção “The public editor” apresenta as opiniões e conclusões dos leitores do jornal em coluna publicada no mínimo duas vezes por mês, aos domingos. Já no *Washingtonpost.com*, o material produzido por estudantes de universidades americanas também pode aparecer incorporado à oferta informativa, assim como as opiniões e comentários dos usuários, desde que registrados no *site*. E o *MSNBC*, por sua vez, publica o conteúdo produzido pelos usuários na subseção “First Person”, dentro da seção intitulada “Blogs”.

Entre os agregadores, o *Topix.net* é o que concede protagonismo aos usuários, publicando os seus comentários sob a denominação “Recent Comments” na *front page*, além de disponibilizar fóruns e as “Top Stories” da blogosfera na seção *blogs*. Como os produtos digitais mais automatizados, uma crítica que pode ser feita aos agregadores é

justamente a de não ativar áreas para a participação dos usuários, o que certamente poderia agregar valor aos mesmos.

O *Libération* tem ampliado continuamente a visibilidade para os comentários e reações dos usuários, além de outras opções de interatividade, mas certamente é a *BBC News* que mais se destaca nos propósitos de envolver o seu público. Cooperou para isso as ações desenvolvidas desde 2003 pela *British Broadcasting Corporation (BBC)*, quando foi lançado o projeto *iCan*, com o objetivo de equipar o público com ferramentas do ativismo político. Assim, além dos espaços (“Your Stories”, “Have your Say” e “On this Day”) para publicar o material dos usuários (gravações em áudio, vídeos, textos, comentários, testemunhos sobre algum evento passado, etc.), a *BBC News* também coloca à disposição *softwares*, entre outras ferramentas para facilitar a confecção das peças. Vale lembrar que, de acordo com a linha editorial do produto, foram estabelecidos os critérios para o público participar, e definidas formas distintas para realizar a moderação nos fóruns (*Fully Moderated* e *Reactively Moderated*), como mencionado no capítulo 3.

A integração dos usuários como fonte informativa requer a delimitação clara de regras e princípios, de acordo com as legislações vigentes, os direitos de autor, a privacidade, e com as linhas editoriais dos respectivos produtos. Até porque, segundo as próprias definições para a autoria nos ambientes digitais, devem-se escrever as regras pelas quais os textos aparecem tanto quanto escrever os próprios textos e, assim, estabelecer regras para o envolver os usuários como navegadores, protagonistas, exploradores e construtores (MURRAY, 1997, 1999, 2003).

Neste sentido, níveis de participação podem ser instituídos para ordenar os fluxos e garantir o bom funcionamento dos sistemas de produção como um todo. Apresentamos no capítulo 1 os quatro níveis sugeridos por Maria Angeles Cabrera

(2005, p.332) para a atuação dos usuários: protagonista, testemunha, colaborador, comentarista, e que podem ser aplicados para auxiliar na criação das seções. Em complementação aos níveis, a autora indica também os modos como os usuários participam: fonte desencadeante, quando parte da própria fonte a iniciativa de difundir um fato concreto; e fonte complementar, quando completa uma informação publicada por iniciativa dos *sites* ou outros meios.

Quanto à sua implicação com os meios, os usuários podem ser fontes para outros produtos ou fontes que são meio (CABRERA, 2005, p.333), no caso de *sites* como *Indymedia Center*, *Slashdot*, *Digg*, *blogs*, *Wikinews*, *Ohmynews*, dentre outros nos quais se trabalha com a concepção de jornalismo participativo (HALL, 2001, 2005; PRYOR, 2002; BOWMAN & WILLIS, 2003; GILLMOR, 2004, 2005; ALVES, 2005; PAVLIK, 2005; SALAVERRÍA, 2005a, 2005c), jornalismo cidadão (ROSEN, 2003, 2005a, 2005b, 2006a, 2006b) ou jornalismo 3.0 (VARELLA, 2005).

Uma vez que os jornalistas passam a dividir, a partilhar o poder com os usuários, estes, por sua vez, também precisam agregar o sentido de responsabilidade por aquilo que criam e submetem aos *sites* jornalísticos, e demais produtos. Além disso, devem ter a compreensão clara de que a atuação como produtores não os convertem em jornalistas.

## **CONCLUSÕES**

## Conclusões

---

*Não há, provavelmente, duas coisas mais difíceis de prever neste mundo do que o futuro da arte e o futuro do software. Essas visões do porvir só podem ser especulações, extrapolações do ambiente atual, que continua mudando mesmo enquanto escrevo.*  
- Janet H. Murray.

Esta tese foi desenvolvida num momento em que a convergência tecnológica de matriz digital se apresenta consolidada como um traço marcante da sociedade contemporânea. Para o jornalismo, a convergência significa integração entre meios distintos, produção de conteúdos combinando multi-plataformas para publicação, convergência estrutural com a reorganização das redações e a introdução de novas funções para os jornalistas, uso associado de tecnologias da informação, *softwares*, e sistemas inteligentes, exploração do potencial interativo, hipertextual e multimídia da internet, e também a construção de narrativas jornalísticas em conformidade com tais recursos.

Neste cenário, as bases de dados foram denominadas como uma nova forma cultural típica do contemporâneo para conformar a produção material que tem o computador como meio para a sua criação e disseminação. Atrelada à essa noção foi que se percebeu o potencial daquela tecnologia da informação para gerar *web sites* jornalísticos mais verdadeiramente originais, porque afinados com parâmetros e características específicas dos ambientes digitais.

A partir daí se delimitou o objetivo geral desta investigação: identificar as remediações, as rupturas e as potencialidades que a crescente utilização das bases de dados gera para a modalidade do jornalismo digital. Desde o início, intuíamos que rupturas qualitativas seriam verificadas - não apenas em relação às modalidades tradicionais, mas, particularmente, mudanças expressivas com os padrões prévios para os *sites* jornalísticos.

Por outro lado, a análise do histórico de adoção das BDs no jornalismo, que se apresentava como uma tarefa essencial e constituiu um dos objetivos específicos, nos permitiu reconhecer remediações ou continuidades e, com elas, identificar os aperfeiçoamentos ocorridos e enxergar as potencialidades. Ao lado disso, pudemos comprovar que a flexibilidade e, sobretudo, o potencial de reinvenção das bases de dados, as recolocavam como protagonistas a favorecer inovação continuada para o campo do jornalismo. Pôde-se confirmar, ainda, que o emprego delas se encontrava num nível ascendente por conta do conjunto de vantagens para a construção de produtos jornalísticos digitais dinâmicos.

Verificou-se, sobretudo, que as bases de dados já não desempenhavam apenas uma função documental e auxiliar como até então eram percebidas, desde o início da sua incorporação ao trabalho jornalístico. Hoje, as bases de dados têm o caráter estruturante e possuem diversas funções – 18 delas foram identificadas nesta investigação – que condicionam os processos e práticas no jornalismo digital.

Ao examinar como as BDs faziam emergir uma forma cultural (outro objetivo específico), detectamos que para o jornalismo digital elas também configuravam um formato e, por conseguinte, alcançavam um novo *status*. O esforço empreendido para a construção de uma fundamentação teórica que nos possibilitasse cumprir os objetivos traçados e confirmar as hipóteses, com a devida articulação com o estudo empírico dos

22 casos analisados, levou-nos à identificação de um modelo distintivo – o do **Jornalismo Digital em Base de Dados (JDBD)**. Tal paradigma foi proposto como designativo de uma fase de transição entre a terceira e uma quarta etapa de desenvolvimento para esta modalidade, pois o cenário atual indica que o próximo estágio está ainda em gestação.

Neste sentido, o capítulo 5 deste trabalho funciona, no contexto do texto produzido, como um “espaço de conclusões”, uma vez que sintetiza o que consideramos sejam as conseqüências centrais da crescente adoção das bases de dados como estruturantes da atividade jornalística em todas as suas dimensões (pré-produção, produção, pós-produção). Cabe ressaltar que as hipóteses – central e secundárias – guiaram a organização dos capítulos da tese, conferindo coerência interna para o tratamento do tema.

Como o processo de migração das operações digitais para produtos estruturados em bases de dados é gradual, o **JDBD** vai se consolidando igualmente de maneira progressiva. De todo modo, seja qual for a configuração final para o estabelecimento da quarta geração de desenvolvimento do jornalismo digital, as BDs e o modelo que ajudam a forjar desempenharão papel fundamental. Foram elas que habilitaram a passagem de um padrão estático para outro dinâmico, fazendo com que se descobrisse o valor da informação estruturada, organizada, para se publicar conteúdos originais.

Com isso, a gestão da informação – e do conhecimento – tornou-se ágil, uma vez que elas permitem a automatização dos processos a partir dos sistemas publicadores, e desvinculam o conteúdo da forma. Assim, facilitam para as operações a tarefa de conceber usos diferentes para os conteúdos, dando a eles destinações diversas: celulares, agendas eletrônicas, *MP3 players*, *iPods*, distribuição do material para outros *web sites*,

*blogs*, e meios sociais. Ou seja, não apenas para o seu produto específico, como ocorria anteriormente com as páginas estáticas.

As funcionalidades e categorias do **JDBD** evidenciam as potencialidades de aplicação das bases de dados associadas aos algoritmos, que são o fundamento para a criação de tecnologias diferenciais para dar forma aos produtos, em conjunto com o uso de linguagens de programação, metadados, *data mining*, técnicas e métodos de visualização para as informações, bem como o uso de recursos para se desenvolver, publicar, distribuir e comercializar os conteúdos.

Quanto à narrativa jornalística em si e aos elementos conceituais que organizam a oferta informativa nos produtos, o modelo proposto propicia igualmente renovação e criatividade. Seguindo os elementos que instituem a hipernarrativa, as propriedades do jornalismo digital e, em particular, aqueles específicos relacionados ao emprego das BDs, outros parâmetros foram apontados para a construção das peças informativas. Como resultado, tem-se uma possibilidade maior para a classificação externa dos conteúdos, também proporcionada pelas co-relações e a hiperlinkagem do material realizada no processo de indexação e classificação interna dos mesmos. Ademais, novas tematizações são verificadas, auxiliando na expansão do eixo temático tradicionalmente trabalhado no jornalismo, que concorrem para uma maior densidade informativa.

Verificou-se, ainda, que a participação dos usuários como produtores de conteúdos se consolida como mais efetiva, obrigando a um reposicionamento de estratégias por parte dos produtos convencionais e daqueles originalmente digitais para envolver o público como protagonista. Por outro lado, detectou-se a necessidade para que os jornalistas agreguem habilidades cognitivas para atuarem nos ambientes inteligentes, nos quais também desempenharão novas funções ao lado daquelas clássicas que consolidaram o jornalismo como um campo específico, e como uma forma singular

de conhecimento da realidade<sup>237</sup>. Ao lado disso, indicou-se a necessidade de as organizações jornalísticas assimilarem um conjunto de saberes e práticas para que possam operar mais satisfatoriamente num contexto de sociedade cada vez mais complexa.

Para além dessas constatações que a realização da tese nos facultou apresentar, confirmando o cumprimento dos objetivos e das hipóteses da investigação, apontamos também ser o uso das bases de dados o aspecto a conduzir o surgimento de uma nova metáfora para a apresentação das informações jornalísticas. Ancorada na *database aesthetics*, a por nós denominada metáfora JDBD se anuncia não apenas como potencialidade, mas concretamente já se materializa em muitos dos produtos que foram analisados. E, portanto, traz uma mudança significativa indicando novos padrões para se visualizar as informações, seja de modo individualizado, seja na configuração do produto como um todo. O emprego desta metáfora como uma estratégia comunicacional específica para produtos digitais poderá ser amplificado posteriormente, posto que o desenvolvimento de *softwares* e as implementações tecnológicas são crescentes.

Podemos concluir que o modelo proposto aporta relevantes contribuições para as configurações futuras do jornalismo digital, conquanto possa conter lacunas e discrepâncias. Pois, um estudo circunscrito ao âmbito dos produtos, embora contemple uma importante esfera e permita observar a concretização dos processos para forjá-los,

---

<sup>237</sup> O conceito de jornalismo como uma forma singular de conhecimento da realidade e diferenciada do conhecimento do senso comum, da arte e da ciência está presente em Adelmo Genro Filho (1987). Robert Park, ex-jornalista e sociólogo norte-americano fundador da sociologia urbana, publicou, em 1940, o artigo *News as a form of knowledge*, no qual define o jornalismo como forma de conhecimento da realidade a partir do que ele tem de diferente e do que lhe é específico. Ele propõe a existência de uma gradação entre um “conhecimento de” utilizado no cotidiano e um “conhecimento sobre”, sistemático e analítico, como o produzido pelas ciências, observando que o jornalismo realiza para o público as mesmas funções que a percepção realiza para os indivíduos (PARK, 1955).

naturalmente indica que se também fossem observadas as rotinas produtivas, certamente traria resultados mais amplos, nos habilitando a demonstrar um alcance maior para eles.

Contudo, da análise do *corpus* empírico, constituído por 22 produtos de perfis distintos, seleccionados pela originalidade, representatividade e diversidade, inferimos a constituição de um cenário no qual as bases de dados são centrais, pois:

- Asseguram o padrão dinâmico para a oferta efetiva de conteúdos originais;
- Habilitam a organização e a estruturação das informações jornalísticas, permitindo a automatização dos processos;
- Conferem maior agilidade para o tratamento das informações nas etapas de pré-produção, produção e pós-produção, assegurando rapidez para a recuperação e para o compartilhamento dos conteúdos;
- Estabelecem-se como soluções essenciais para a gestão do conhecimento nas redações cada vez mais integradas e operando segundo a lógica de um jornalismo convergente;
- Favorecem a expansão de atuação para as empresas, pois permitem a distribuição do conteúdo jornalístico para plataformas variadas;
- Garantem renovação continuada para os conteúdos, pois são uma tecnologia flexível que pode estar associada a algoritmos, linguagens de programação, *softwares*, metadados, *data mining*, entre outras técnicas;
- Proporcionam experimentações com novos padrões estéticos para os *sites* jornalísticos, permitindo romper com modelos transpositivos prévios, sobrepujar a metáfora do impresso e demarcar uma estratégia de comunicação própria;

- Propiciam a ampliação do eixo temático e a conseqüente diversidade temática, através da agregação do material de arquivo na oferta informativa, do conteúdo produzido pelos usuários, e do inter-relacionamento entre as informações produzidas pelo próprio *site*, assim como aquelas provenientes de meios externos.

### **Desdobramentos**

Um trabalho doutoral não se restringe aos quatro anos de sua elaboração, impulsionando o pesquisador a dar continuidade à investigação, explorando vertentes novas e verificando os desdobramentos que possam advir da realização dela. Sendo assim, colocamos como continuidade do estudo a realização do pós-doutorado, durante o qual se trabalhará as funcionalidades e categorias do modelo **JDBD** aplicado às rotinas produtivas em *web sites* metainformativos, inteligentes e dinâmicos de dois países (Brasil e Espanha) e com a sua inserção no âmbito de um projeto de pesquisa maior que congrega investigadores de universidades nacionais e espanholas. O estudo comparativo permitirá conhecer os processos internos de produção, o nível de utilização das BDs na gestão interna das empresas e dos respectivos produtos, avaliar a destreza dos profissionais ao operar com bases de dados e os sistemas de gestão de conteúdos, bem como verificar algumas conseqüências da incorporação de BDs no processo de apuração, circulação e comercialização de informações.

Esta tese, portanto, reflete o esforço realizado para compreender os usos e as transformações que a crescente utilização das bases de dados está fazendo emergir para o jornalismo digital. Como tal, consideramos que o trabalho desenvolvido traz não apenas a cartografia de um cenário, ou a detecção de uma tendência. Reconhecemos a

sua importância por ter focalizado um aspecto singular, que, como confirma o processo histórico, tem exercido papel de grande relevância para o jornalismo e, para o jornalismo digital, se apresenta como a especificidade que diferencia esta modalidade das tradicionais.

Neste sentido, ressalta-se a pertinência do estudo para a área, pois representa um empenho inicial de sistematização, de definição e de delimitação de um objeto novo, que é a apropriação das bases de dados digitais pelo jornalismo. Assim, a tese concretiza uma pesquisa inédita no país, com a proposição do modelo **Jornalismo Digital em Base de Dados**, o qual certamente integrará as próximas etapas de evolução do jornalismo digital, já que as BDs são elementos estruturantes e como tal condicionam a criação de produtos jornalísticos digitais.

Por fim, gostaríamos de registrar que se a tese reflete um cariz original e inovador isso se deve ao contexto da qual ela emergiu: o Grupo de Pesquisa em Jornalismo On-line da FACOM/UFBA que integra o Centro Internacional de Estudos e Pesquisa em Cibercultura (Ciberpesquisa). Foi lá que nos últimos quatro anos – além do período de realização do mestrado – moldamos a nossa formação acadêmica, guiada por professores de extrema competência e espírito provocador sempre a nos estimular. Esperamos que o trabalho gerado esteja à altura.

\*\*\*

- Janeiro/Fevereiro de 2007. Entre o inverno lusitano e o verão soteropolitano.

---

## Referências bibliográficas

---

### REFERÊNCIAS BIBLIOGRÁFICAS

#### Livros, artigos e periódicos correntes (meio impresso e meio digital).

- ADAMS, Katherine C. **The Web as Database: New Extraction Technologies and Content Management**. Março, 2001. Disponível em: [http://www.infotoday.com/Online/OL2001/adams3\\_01.html](http://www.infotoday.com/Online/OL2001/adams3_01.html). Acesso em: 28/11/2005.
- AGUILERA, Isabel. **La experiencia de ser googly**. 2006. In: *Revista Telos – Cuadernos de comunicación, tecnología y sociedad*. Outubro-Diciembre 2006, nº29. Disponível em: <http://www.campusred.net/telos/articulocuaderno.asp?idArticulo=9&rev=69>. Acesso em: 08/11/2006.
- ÄKESSON, Maria. **Design patterns for the online newspaper genre: a genre analysis of 85 swedish daily online newspapers**. Junho, 2003. Disponível em: [www.handels.gu.se/epc/archive/00002936/01/Nr10\\_M%C3%85.pdf](http://www.handels.gu.se/epc/archive/00002936/01/Nr10_M%C3%85.pdf). Acesso em: 01/03/2004.
- ALVES, Rosental. **Jornalismo digital: Dez anos de Web...e a revolução continua**. Conferência proferida nas jornadas “Dez Anos de Jornalismo Digital em Portugal – Estado da Arte e Cenários Futuros” ([www.dezanos.blogspot.com](http://www.dezanos.blogspot.com)), organizadas pelo projecto Mediascópio, do Centro de Estudos de Comunicação e Sociedade da Universidade do Minho, Braga (Portugal), 02 de Junho, 2005.
- \_\_\_\_\_. **Reinventando o jornal na internet - A internet: chegou a hora de reavaliar sua função no jornal impresso**. [mensagem pessoal]. Mensagem recebida por [suzana-b@uol.com.br](mailto:suzana-b@uol.com.br) em 25/03/2001.
- ALVES, Maria Bernardete M.; ARRUDA, Susana Margareth. **Como fazer referências** (bibliográficas, eletrônicas e demais formas de documentos). 2000/2003. Disponível em: <http://www.bu.ufsc.br/framerefer.html>. Acesso em: 02/09/2006.

- ALZAMORA, Geane Carvalho. **Webjornalismo/Semiótica/Informação**. Disponível em: <http://www.razonypalabra.org.mx/anteriores/n32/gcarvalho.html>. Acesso em: 06/06/2005.
- ARMAÑANZAS, E.; NOCI, J.D.; MESO, K. **El periodismo electrónico**. Información y servicios multimedia en la era del ciberespacio. Barcelona, Ariel Comunicación, 1996.
- ARNS, Ink. **Metonymical Mov(i)es**. 2002. Disponível em: <http://www.projects.v2.nl/~arns/Texts/Media/manovich-review.html>. Acesso em: 10/09/2006.
- ATTWOOD, Gaynor; MAIDENHEAD, Ian Attwood. **Information Technology Dictionary**. Berkshire: England, 1989.
- BALDESSAR, Maria José. **A mudança anunciada**. O cotidiano dos jornalistas com o computador na redação. Florianópolis: Insular, 2003.
- BARBOSA, Suzana. **O que é jornalismo digital em bases de dados**. In: Anais XV Encontro Anual da Compós (CD-ROM). Bauru: Editora UNESP, 2006a., v. 1.
- \_\_\_\_\_. **Jornalismo digital em bases de dados**. In: VII Congreso Internacional de Comunicación Lusófona – LUSOCOM. Actas do Congresso Internacional Lusocom 2006. Santiago de Compostela: Servizo de Publicacións e Intercambio Científico - Universidade de Santiago de Compostela, 2006b., v. 168.
- \_\_\_\_\_. **Sistematizando conceitos e características sobre o Jornalismo Digital em Base de Dados**. Comunicação apresentada nas Jornadas Jornalismo Online.2005. Aspectos e Tendências. Universidade da Beira Interior, Portugal, Novembro, 2005c.
- \_\_\_\_\_. **Bases de dados e webjornalismo: em busca de novos conceitos**. In: 4º SOPCOM - Repensar os Media: Novos Contextos, 2005, Aveiro. 4º SOPCOM - Livro de Actas do 4º Congresso da Associação Portuguesa de Ciências da Comunicação (Mesa Novas Tecnologias. Novas Linguagens). Aveiro: Universidade de Aveiro, 2005b.
- \_\_\_\_\_. **Jornalismo digital e bases de dados: mapeando conceitos e funcionalidades**. In: 4º SOPCOM - Repensar os Media: Novos Contextos, 2005, Aveiro. 4º SOPCOM - Livro de Actas do 4º Congresso da Associação Portuguesa de Ciências da Comunicação (Mesa Jornalismo). Aveiro: Universidade de Aveiro, 2005a.
- \_\_\_\_\_. **Bancos de Dados: agentes para um webjornalismo inteligente?** In: V Congreso Iberoamericano de Periodismo en Internet, 2004, Salvador. V Congreso Iberoamericano de Periodismo en Internet. Salvador: Agecom, 2004c.
- \_\_\_\_\_. **Identificando remediações e rupturas no uso de bancos de dados no jornalismo digital**. In: Anais do II SBPJor. Salvador-BA/Brasil, 2004b.

- \_\_\_\_\_. **Banco de Dados como metáfora para o jornalismo digital de terceira geração.** In: Anais III Sopcom, VI Lusocom, II Ibérico, UBI (CD-ROM), 2004a.
- \_\_\_\_\_. **Jornalismo de portal: novo formato e categoria para o jornalismo digital.** In: MACHADO, E.; PALACIOS, M. (Orgs.). **Modelos de jornalismo digital.** Salvador: Calandra, 2003.
- \_\_\_\_\_. **Jornalismo digital e a informação de proximidade: o caso dos portais regionais, com estudo sobre o UAI e o iBAHIA (Dissertação de Mestrado).** FACOM-UFBA, Salvador, 2002.
- BARBOSA, Suzana; MIELNICZUK, Luciana; QUADROS, Claudia. **Estudos sobre jornalismo digital no Brasil.** In: *eCompós*. Revista da Associação Nacional dos Programas de Pós-Graduação em Comunicação. Edição nº7, Dezembro, 2006. Disponível em: [http://www.compos.org.br/ecompos/adm/documentos/ecompos07\\_dezembro2006\\_claudia\\_luciana\\_suzana.pdf](http://www.compos.org.br/ecompos/adm/documentos/ecompos07_dezembro2006_claudia_luciana_suzana.pdf). Acesso em: 23/01/2007.
- \_\_\_\_\_. **A Study about Digital Journalism in Brazil.** 2006. In: Latin American Network Center (LANIC), University of Texas at Austin, EUA. Disponível em: <http://lanic.utexas.edu/project/etext/>. Acesso em: 20/09/2006.
- BARBOSA, Suzana; MACHADO, Elias; MIELNICZUK, Luciana; PALACIOS, Marcos; RIBAS, Beatriz. **Beyond Prophecies: A critique of “New News Retrospective”.** In: *Brazilian Journalism Research. Journalism theory, research and criticism*. Vol.2, Nº.1 (jan./jun.2006). Brasília: Universidade de Brasília. SBPJor, Brazilian Association for Journalism Researchers.
- BARDOEL, Jo & DEUZE, Mark. **Network journalism: converging competences of old and new media professionals.** In: <http://home.pscw.uva.nl/deuze/pub19.htm>. Acesso em: 13/10/2000.
- BASTOS, Hélder. **Jornalismo Electrónico.** Internet e Reconfiguração de Práticas nas Redacções. Coimbra: Minerva, 2000.
- BELL, Allan. **The language of news media.** Cambridge, Massachusetts: Blackwell Publishers, 1991.
- BELL, David et all. **Cyberculture.** The key concepts. London: Routledge, 2004.
- BENÍTEZ, A. S.; RODRÍGUEZ, Antonio A.R. **Metadatos para la preservación de colecciones digitales.** In: Cuadernos de Documentación Multimedia, V.16, 2005. Disponível em: <http://multidoc.rediris.es/cdm/index.php?OJSSESSID=341be5359ef8b1de8b12f22e2b9b3985> . Acesso em: 08/09/2005.
- BERLINSKI, David. **O advento do algoritmo: a idéia que governa o mundo.** São Paulo: Globo, 2002.

- BJORNER, Susanne; ARDITO, Stephanie C. [Online Before the Internet: Early Pioneers Tell Their Stories](#). Part 1: In the Beginning. 2003. Disponível em: <http://www.dialog.com/about/history/pioneers1.pdf>. Acesso em: 27/04/2005.
- BOCZKOWSKI, Pablo J. **Digitizing The News**. Innovation in Online Newspapers. The MIT Press, 2004.
- \_\_\_\_\_. **The development and use of online newspapers: what research tell us and what we might want to know**. In: LIEVROUW, Leah A.; LIVINGSTONE, Sonia. **Handbook of new media**. Social Shaping and Consequences of ICTs. London: Sage, 2002.
- BOLTER, Jay David; GROMALA, Diane. **Windows and Mirrors**. Interaction design, digital art, and the myth of transparency. Cambridge: The MIT Press, 2003.
- BOLTER, Jay David; GRUSIN, Richard. **Remediation: Understanding new media**. Cambridge: The MIT Press, 2000.
- BOLTER, Jay David. **Writing Space**. The computer, hypertext and the history of writing. Hilldale: LEA, 1991.
- BORDEN, Diane L.; HARVEY, Kerric. **The electronic grapevine**. Rumor, Reputation, and Reporting in the New On-line Environment. Mahwah, New Jersey, London: LEA, 1998.
- BOULTER, Jeff. **Online publishing: The Past, Present and Future of Electronic Distribution**. 1995. Disponível em: <http://www.nepressence.com/boulter/jeff/onlinepub.html>. Acesso em: 02/05/2006.
- BOWERS, Greg. **Web for 'News' - Print for 'Stories'**. In: Editor & Publisher, julho, 2006. Disponível em: [http://www.editorandpublisher.com/eandp/columns/shoptalk\\_display.jsp?vnu\\_content\\_id=1002876835](http://www.editorandpublisher.com/eandp/columns/shoptalk_display.jsp?vnu_content_id=1002876835). Acesso em: 23/07/2006.
- BOWMAN, Shayne; WILLIS, Chris. **We media**. How audiences are shaping the future of news and information. Edited by J.D.Lasica para The Media Center at The American Press Institute (Julho/2003). Disponível em: [www.hypergene.net/wemedia](http://www.hypergene.net/wemedia); [www.mediacenter.org/mediacenter/research/wemedia](http://www.mediacenter.org/mediacenter/research/wemedia). Acesso em: 23/03/2005.
- BRESSAN, Flavio. **O método do estudo de caso**. Disponível em: [http://www.fecap.br/adm\\_online/art11/flavio.htm](http://www.fecap.br/adm_online/art11/flavio.htm). Acesso em: 10/03/2001.
- BRETON, Philippe; PROULX, Serge. **A explosão da comunicação**. Lisboa: Bizâncio, 2000.
- BRICKLIN, Dan. **The Cornucopia of the Commons: How to get volunteer labor**. 2001. Disponível em: <http://www.bricklin.com/cornucopia.htm>. Acesso em: 29/08/2006.

- BROWN, Edward y CHIGNELL, Mark H. **El usuario como diseñador: el multimedia de forma abierta**. In: **Medios contextuales en la práctica cultural**. Barcelona: Piados, 1997.
- BURKE, Peter & BRIGGS, Asa. **Uma história social da mídia**. De Gutenberg à Internet. Rio de Janeiro: Jorge Zahar Editor, 2004.
- CABRERA, Maria Angeles. **Retos éticos del ciberperiodismo**. In: SALAVERRÍA, Ramón. (Coord.). [Cibermedios. El impacto de internet en los medios de comunicación en España](#). Sevilla: Comunicación Social Ediciones y Publicaciones, 2005. p: 305-339.
- CANAVILHAS, João. **Webjornalismo: Da pirâmide invertida à pirâmide deitada**. Comunicação apresentada nas Jornadas Jornalismo Online.2005. Aspectos e Tendências. Universidade da Beira Interior, Portugal, Novembro, 2005.
- \_\_\_\_\_. **Webperiodismo: de la pirâmide invertida a la pirâmide tumbada**. In: VII Congreso Internacional de Comunicación Lusófona – LUSOCOM. Actas do Congreso Internacional Lusocom 2006. Santiago de Compostela: Servizo de Publicacións e Intercambio Científico - Universidade de Santiago de Compostela, 2006, v. 168.
- \_\_\_\_\_. **A internet como memória**. (Fevereiro/2004). Disponível em: <http://www.webjornalismo.com/sctions.php?op=viewarticle&artid=78>. Acesso em: 07/03/2004.
- \_\_\_\_\_. **Webjornalismo: considerações gerais sobre jornalismo na web**. 2001. Disponível em: [http://www.bocc.ubi.pt/pag/\\_texto.php3?html2=canavilhas-joao-webjornal.html](http://www.bocc.ubi.pt/pag/_texto.php3?html2=canavilhas-joao-webjornal.html). Acesso em: 25/03/2002.
- CARA, Francesco; BROADBENT, Stephana. **Les nouvelles architectures de l'information**. In : ORIGGI, Gloria; ARIKHA, Noga. **Le texte à l'heure de l'internet**. Paris: Bibliothèque publique d'information, 2003. pp. 197-213.
- CARDOSO, C. G.; ARAÚJO, João Gualberto Rizzo de ; SAMPAIO, M. A. **Uma Breve História da Internet na Bahia**. In: *Revista da Universidade de Aveiro*, v. 1, n. 1, p. 54-67, 2002, Aveiro, Portugal.
- CARDOSO, C. G. **Internet com Ferramenta para Negócios**. 2001. In: *404NOTFOUND*. Disponível em: [www.404notfound.com.br](http://www.404notfound.com.br). Acesso em: 26/05/2005.
- CARITÁ, Enrico. **Já estamos no futuro**. In: **Evolução na comunicação: do sílex ao silício**. Rio de Janeiro: Nova Fronteira, 1987. pp: 280-374.
- CARLEY, Kathleen M. **Smart Agents and Organizations of the Future**. In: LIEVROUW, Leah A.; LIVINGSTONE, Sonia. **Handbook of new media**. Social Shaping and Consequences of ICTs. London: Sage, 2002.

CASTELLS, Manuel. **La Galáxia Internet**. Barcelona: Areté, 2001.

\_\_\_\_\_. **A era da informação: economia, sociedade e cultura**. Vol.1: a sociedade em rede. São Paulo: Paz e Terra, 1999.

CASTILHO, Carlos. **Jornalismo online**. Cada cidadão é um repórter. Disponível em: [www.observatoriodaimprensa.com.br](http://www.observatoriodaimprensa.com.br), edição 275 de 04/05/2004. Acesso em: 07/05/2004.

CENDÓN, Beatriz Valadares. **Bases de dados de informação para negócios**. 2002. Disponível em: <http://www.scielo.br/pdf/ci/v31n2/12906.pdf>. Acesso em: 25/10/2006.

CHEN, Chaomei. **Information visualization: Beyond the horizon**. (Second edition). London: Springer-Verlag, 2004.

CISLER, Steve. **Arts, crafts, and globalization**. Databases of art and crafts will be more complete and in greater richness than was possible. 2000. Disponível em: [http://switch.sjsu.edu/nextswitch/switch\\_engine/front/front/php?artc=11](http://switch.sjsu.edu/nextswitch/switch_engine/front/front/php?artc=11). Acesso em: 28/07/2006.

CODINA, Lluís; RECODER, Maria-José; ABADAL, Ernest. **Informação eletrônica e novas tecnologias**. Tradução de Dinah Aguiar Población. São Paulo: Summus Editorial, 1995. (Coleção Novas Buscas em Comunicação).

COLLE, Raymond. **Explotar la información noticiosa**. *Data mining* aplicado a la documentación periodística. Madrid: Departamento de Biblioteconomía y Documentación, Universidad Complutense de Madrid, 2002.

\_\_\_\_\_. **5000 Notícias sobre tecnologías digitales de comunicación**. *TDC*, 2002/2003. Disponível em: <http://www2.ing.puc.cl/~dcolle/publicaciones/5000/5000not.htm> e também em: <http://www2.ing.puc.cl/~dcolle/publicaciones/5000not.pdf>. Acesso em: 07/08/2005.

\_\_\_\_\_. **Processos documentales y gestión del conocimiento**. 2005b. In: *Razón y Palabra*, N°4, ago/sept. Disponível em: [www.razonypalabra.org.mx/actual/rcolle.html](http://www.razonypalabra.org.mx/actual/rcolle.html). Acesso em: 06/09/2005

\_\_\_\_\_. **Metainformación**. La ventaja inexplorada de la comunicación on-line. 2005a. Disponível em: [www2.ing.puc.cl/~dcolle/publicaciones/metainfo/meta.htm](http://www2.ing.puc.cl/~dcolle/publicaciones/metainfo/meta.htm). Acesso em: 06/09/2005.

\_\_\_\_\_. **Análisis lógico de hechos noticiosos**. 2000. In: *Revista Latina de Comunicación Social*, N°37. Disponível em: [www.ull.es/publicaciones/latina](http://www.ull.es/publicaciones/latina). Acesso em: 13/09/2005.

COLOMBO, Fausto. **Os arquivos imperfeitos**. São Paulo: Perspectiva, 1991.

**COMUNICAÇÃO DIGITAL: COMPETÊNCIAS PROFISSIONAIS E DESAFIOS ACADÉMICOS**. REDE ICOD. Primeira edição: Abril de 2006.  
**CONGRESSO INTERNACIONAL E SEMINÁRIO A CULTURA DAS REDES EM PORTUGAL. REPRESENTATION IN THE 21 CENTURY: MEDIA,**

**NEW MEDIA, AND META-MEDIA** (Palestra com Lev Manovich). Centro Cultural Belém. Lisboa, Portugal, 26 e 27 de Maio, 2006.

CRESPO, Ángela M.<sup>a</sup> Cavalcanti; SAGREDO, Félix. **La visión de Jerome Aumente sobre la alta tecnología de la edición electrónica (tres aportaciones: 1988-1994)**. Disponível em:

<http://www.ucm.es/BUCM/revistas/inf/02104210/articulos/DCIN9696110043A.PDF>. Acesso em: 30/07/2006.

DARNTON, Robert. **The new age of the book**. In: *The New York Review of Books*. 18/março, 1999; pp.5-7. Vol.46, N.5. Disponível em: [www.nybooks.com/articles/546](http://www.nybooks.com/articles/546). Acesso em: 10/05/2006.

DATE, C.J. **Introdução a sistemas de bancos de dados**. (Título original: *An introduction to database systems*). Rio de Janeiro: Editora Campus, 1984.

DEEN, S.M. **Fundamentals of data base systems**. London and Basingstoke: The Macmillan Press LTD, 1977.

DEUZE, Mark. **What is Multimedia Journalism?** In: *Journalism Studies*, vol. 5, n.º. 2, pp. 139-52, 2004. Disponível em:

<http://convergence.journalism.indiana.edu/media/documents/convergence/DeuzeMultimediaJS.pdf>. Acesso em: 31/08/2005.

\_\_\_\_\_. **The web and its journalisms: considering the consequences of different types of newsmedia online**. In: *New Media & Society*. London, Thousand Oaks, CA and New Delhi: Sage Publications, 2003.

DEWDNEY, Andrew y BOYD, Frank. **La television, los ordenadores, la tecnología y la forma cultural**. In: LISTER, Martin (Org.). **La imagen fotográfica en la cultura digital**. Barcelona: Paidós, 1997.

DE WOLK, Roland. **Online Journalism**. Publishing news and information. Boston: Allyn and Bacon, 2001.

DIAZ NOCI, Javier. **La escritura digital**. Hipertexto y construcción del discurso informativo en el periodismo electrónico. Bilbao: Universidad del País Vasco, 2001.

DIAZ NOCI, Javier; SALAVERRÍA, Ramón. **Manual de Redación Ciberperiodística**. Barcelona: Ariel, 2003.

DÍAZ NOCI, Javier; AYERDI, Koldo Meso. **Periodismo en Internet: modelos de la prensa digital**. Bilbao: Universidad del País Vasco, 1999.

DIETZ, Steve. **The Database Imaginary: Memory\_Archive\_Database v 4.0**. (2004). Disponível em: <http://databaseimaginary.banff.org/getEssay.php?id=1&t=1>. Acesso em: 28/07/2006.

\_\_\_\_\_. **Memory\_Archive\_Database** V.3.0. Database culture is only part a reflection of the rise of the internet. 2000. Disponível em: [http://switch.sjsu.edu/nextswitch/switch\\_engine/front/front.php?artc=31](http://switch.sjsu.edu/nextswitch/switch_engine/front/front.php?artc=31). Acesso em: 28/07/2006.

**DICTIONARY OF INFORMATION TECHNOLOGY.** (Third Edition). London: Peter Collin Publishing, 2002.

DIZARD Jr, Wilson. **A nova mídia.** A comunicação na era da informação. Rio de Janeiro: Jorge Zahar Editores, 2000.

DORNFEST, Rael; BRICKLEY, Dan. **The power of metadata.** 2006. In: O'Reilly P2P. Disponível em: <http://www.openp2p.com/pub/a/p2p/2001/01/18/metadata.html>. Acesso em: 26/07/2006.

DIETRICH, Frank. **Data particles.** Meta-data-dataspace. 2000. Disponível em: [http://switch.sjsu.edu/nextswitch/switch\\_engine/front/front.php?artc=12](http://switch.sjsu.edu/nextswitch/switch_engine/front/front.php?artc=12). Acesso em: 28/07/2006.

DUIN, Aline van; GANGAHAR, Anuj. **At Thomson, computers are making headlines.** In: Financial Times.Com. 18/Agosto/2006. Disponível em: <http://search.ft.com/searchArticle?queryText=Thomson&y=7&javascriptEnabled=true&id=060818007491&x=17>. Acesso em: 20/09/2006

DUIN, Aline van. **Computers write news at Thomson.** In: Financial Times.Com. 17/Agosto/2006. Disponível em: [www.ft.com/cms/s/bb3ac0f6-2e1511db-93ad-0000779e2340.dwp\\_uuid=cbad99](http://www.ft.com/cms/s/bb3ac0f6-2e1511db-93ad-0000779e2340.dwp_uuid=cbad99). Acesso em: 27/08/2006.

ECHEVERRÍA, Javier. **Los señores del aire:** Telépolis y el tercer entorno. Barcelona: Destino, 1999.

\_\_\_\_\_. **Telépolis.** Barcelona: Destino, 1994.

EDO, Concha. **El lenguaje y los géneros periodísticos en la narrativa digital.** Comunicação apresentada nas Jornadas Jornalismo Online.2005. Aspectos e Tendências, Universidade da Beira Interior, Portugal, Novembro, 2005.

\_\_\_\_\_. **Del papel a la pantalla.** La prensa en internet. Sevilla: Comunicación Social Ediciones y publicaciones, 2002 (Coleção Periodística, 2).

ERIKSEN, Lars Bo. & IHLSTRÖM C. **Evolution of the web news genre – The slow move beyond the print metaphor.** In: Proceedings of 33' Hawaii International Conference on Systems Science. IEEE Press, 2000. Disponível em: [www.computer.org/proceedings/hicss/0493/04933/04933014.pdf](http://www.computer.org/proceedings/hicss/0493/04933/04933014.pdf). Acesso em: 01/03/2004.

FABRITIUS, Hannele. **Information Seeking in the Newsroom.** Application of the Cognitive Framework for Analysis of the Work Context. Disponível em: <http://informationr.net/ir/4-2/isic/fabritiu.html>. Acesso em: 09/07/2006.

- FAYYAD, Usama; PIATETSKI-SHAPIRO, Gregory; SMYTH, Padhraic. **The KDD Process for Extracting Useful Knowledge from Volumes of Data**. In: Communications of the ACM, pp.27-34, Nov.1996. Disponível em: <http://delivery.acm.org/10.1145/250000/240464/p27fayyad.pdf?key1=240464&key2=2311094511&coll=ACM&dl=ACM&CFID=3005669&CFTOKEN=9304936>  
1. Acesso em: 06/08/2006
- FERRARI, Pollyana. **Jornalismo digital**. São Paulo: Contexto, 2003.
- FIDALGO, António. **Data Mining e um novo jornalismo de investigação**. Notícias de Ordem Superior. Comunicação apresentada nas Jornadas *Jornalismo Online.2005. Aspectos e Tendências*. Universidade da Beira Interior, Covilhã, Portugal. Novembro de 2005.
- \_\_\_\_\_. **Do poliedro à esfera: os campos de classificação. A resolução semântica no jornalismo online**. In: Anais do II SBPJor. Salvador-BA/Brasil, 2004.
- \_\_\_\_\_. **Sintaxe e semântica das notícias on-line**. Para um jornalismo assente em base de dados. In: FIDALGO, António; SERRA, Paulo (Orgs.). **Informação e Comunicação Online**. Jornalismo Online. Volume 1. Covilhã: Universidade da Beira Interior/Portugal, 2003.
- \_\_\_\_\_. **O realismo da fenomenologia de Munique**. Braga: Publicações da Faculdade de Filosofia, 1991.
- FIDALGO, António; SERRA, Paulo (Orgs.). **Informação e Comunicação Online**. Jornalismo Online. Volume 1. Covilhã: Universidade da Beira Interior/Portugal, 2003.
- FIDLER, Roger. **Mediamorphosis**. Understanding new media. Thousand Oaks, California: Pine Forge Press, 1997.
- FILHO, Adelmo Genro. **O segredo da pirâmide**. Para uma teoria marxista do jornalismo, Porto Alegre, Tchê, 1987.
- FONTCUBERTA, Mar de. **El temario periodístico**. In: FONTCUBERTA, Mar de; BORRAT, Héctor. **Periódicos: sistemas complejos, narradores en interación**. Buenos Aires: La Crujía, 2006. pp.55-87.
- \_\_\_\_\_. **A notícia**. Pistas para compreender o mundo. Lisboa: Editorial Notícias (Colecção Media & Sociedade), 2ª ed., 2002.
- FONTCUBERTA, Mar de; BORRAT, Héctor. **Periódicos: sistemas complejos, narradores en interación**. Buenos Aires: La Crujía, 2006.
- FOUCAULT, Michel. **A arqueologia do saber**. Rio de Janeiro: Forense Universitária, 7ª ed., 2005.

- FRANCISCATO, Carlos Eduardo. **A fabricação do presente**. Como o jornalismo reformulou a experiência do tempo nas sociedades ocidentais. São Cristóvão: Editora UFS; Aracaju: Fundação Oviedo Teixeira, 2005.
- FRANGOMENI, Ana Helena. **Dicionário enciclopédico de informática**. Campus: Rio de Janeiro; Nobel: São Paulo, 1987. Vol. 1.
- GAGO, Manuel. **La arquitectura de la información, ingeniería del periodismo**. In: LÓPEZ, Xosé (Coord.). **Sistemas digitais de información**. Madrid: Pearson Educación, 2006. pp: 81-142.
- GARCÍA, Xosé L. et al. **Tipología de los cibermedios**. In: SALAVERRÍA, Ramón. (Coord.). **Cibermedios. El impacto de internet en los medios de comunicación en España**. Sevilla: Comunicación Social Ediciones y Publicaciones, 2005. p: 39-82.
- GARRISON, Bruce. **Computer-Assisted Reporting**. London: LEA Publishers, 1998. (Second edition).
- GARRISON, Bruce; DUPAGNE, Michel. **The meaning and influence of convergence**. A qualitative case study of newsroom work at the Tampa News Center. Maio/2005. Disponível em: <http://com.miami.edu/car/JStudies06.pdf>. Acesso em: 09/03/2006.
- GENRO FILHO, Adelmo. **O segredo da pirâmide - para uma teoria marxista do jornalismo**. Porto Alegre: Tchê, 1987.
- GILLMOR, Dan. **Nós, Os Media**. Lisboa: Editorial Presença, 2005.
- \_\_\_\_\_. **We The Media: Grassroots Journalism by the People, for the People**. Cambridge: O'Reilly Media Inc., 2004.
- GIL, Antônio Carlos. **Como elaborar projetos de pesquisa**. 4ª ed. São Paulo: Atlas, 2002.
- GIOVANNINI, Giovanni. **Evolução na Comunicação**. Do sílex ao silício. Rio de Janeiro: Nova Fronteira, 1987.
- GOSCIOLA, Vicente. **Roteiro para as Novas Mídias**. Do game à TV interativa. São Paulo: Senac, 2003.
- GRADIM, Anabela. **Webjornalismo e a profissão de jornalista: alguns equívocos sobre a dissolução do quarto poder**. Comunicação apresentada nas Jornadas Jornalismo Online.2005. Aspectos e Tendências, Universidade da Beira Interior, Portugal, Novembro, 2005.
- \_\_\_\_\_. **O jornalista multimedia do século XXI**. In: FIDALGO, António; SERRA, Paulo (Orgs.). **Informação e Comunicação Online**. Jornalismo Online. Volume 1. Covilhã: Universidade da Beira Interior/Portugal, 2003. p: 117-134.

- GRAHAM, Stephen. **Beyond the 'dazzling light'**: from dreams of transcendence to the 'remediation' of urban life. *A research manifesto*. In: *New Media & Society*, Vol.6 (1): 16-25, 2004. Disponível em: [www.sagepublications.com](http://www.sagepublications.com). Acesso em: 18/03/2004.
- GRANIERI, Giuseppe. **Blog generation**. Roma-Bari: Laterza & Figli Spa, 2005.
- GROSSMANN, Fabiane V. **Comunicação e design de interface**: Uma contribuição para a discussão de estratégias comunicacionais do webjornal. (Exame de qualificação). PPGCOM/UFSM, Santa Maria/RS, 2006.
- GRILO, Nuno. **Organização da base de dados do Público Online**. 1999. Disponível em: <http://xldb.fc.ul.pt/ariadne/documentos/rt-99-1.pdf>. Acesso em: 15/05/2006.
- GUIMARÃES, Célio Cardoso. **Fundamentos de bancos de dados**. Modelagem, projeto e linguagem SQL. Campinas (SP): Unicamp, 2003.
- GUNTER, Barrie. **News and the net**. Mahwah, New Jersey, London: LEA, 2003.
- GUTIÉRREZ DE MESA, José A., GONZALEZ, José Ramón. **Generación de documentación hipermedia en internet a partir de información multimedia en Bases de Datos**. Disponível em: Cuadernos de Documentación Multimedia. Nº 6/7, 1997/1998, <http://www.ucm.es/info/multidoc/multidoc/revista/cuad6-7/hilera.htm>. Acesso em: 08/09/2006.
- HALL, Jim. **The News Blog in 2005: Social Journalism at the Eye of the Storm**. Comunicação apresentada nas Jornadas Jornalismo Online.2005. Aspectos e Tendências, Universidade da Beira Interior, Portugal, Novembro, 2005.
- \_\_\_\_\_. **Online journalism**. A critical primer. London: Pluto Press, 2001.
- HARVEY, David. **Condição pós-moderna**. São Paulo: Loyola, 9ª ed., 2000.
- HEARST, Marti A. **Text Data Mining**. In: MITKOV, Ruslan. Oxford Handbook of Computational Linguistics. Oxford: Oxford University, 2003.
- HEINONEN, A.; MAKINEN, S; et alli. **Locality in the Global Net**: the internet as a space of citizen communication and local publicness. University of Tampere research report, 2000. Disponível em: <http://mansetori.uta.fi/report/default.htm>. Acesso em: 17/10/2003.
- HERBERT, John. **Journalism in the digital age**. Theory and Practice for Broadcast, Print and On-line Media. Oxford: Focal Press, 2001.
- HERING, Susan C. **Slouching toward the ordinary**: current trends in computer-mediated communication. In: *New Media & Society*, Vol 6 (1): 26-36, 2004. Disponível em: [www.sagepublications.com](http://www.sagepublications.com). Acesso em 18/03/2004.

- HOLOVATY, Andrew. **A fundamental way newspaper sites need to change.** 06/Setembro/2006b. Disponível em: <http://www.holovaty.com/blog/archive/2006/09/06/0307>. Acesso em: 04/10/2006.
- \_\_\_\_\_. **Dynamic News Stories.** In: O'Reilly XML.com, 17/Maio/2006a. Disponível em: <http://www.xml.com/pub/a/2006/05/17/dynamic-news-stories.html>. Acesso em: 30/06/2006.
- \_\_\_\_\_. **Microformats could describe online news intelligently.** 2005. Disponível em: [www.holovaty.com/blog/archive/2005/03/18/0033](http://www.holovaty.com/blog/archive/2005/03/18/0033). Acesso em: 30/06/2006.
- HOUSTON, Brant. **Computer-Assisted Reporting.** A practical guide. (Third edition). Boston/New York: Bedford St. Martin's, 2004.
- IHLSTRÖM, Carina, & LUNDBERG, Jonas. **A Genre Perspective on Online Newspaper Front Page Design.** In: *Journal of Web Engineering*, 2004, Vol. 3, Nº. 1, pp. 50-74. Também disponível em: <http://csdl.computer.org/comp/proceedings/hicss/2004/2056/04/205640102.pdf>. Acesso em: 14/12/2006.
- IHLSTRÖM, Carina, & ÅKESSON, Maria. **Genre characteristics – Front page analysis of 85 swedish online newspapers.** 2004. Disponível em: <http://csdl.computer.org/comp/proceedings/hicss/2004/2056/04/205640102.pdf>. Acesso em: 01/03/2004.
- JAMBEIRO, Othon. **Pressupostos básicos para a construção de políticas de comunicação na América Latina.** 2006. In: *Verso e reverso revista da comunicação*, Nº44. Disponível em: [www.versoereverso.unisinos.br](http://www.versoereverso.unisinos.br). Acesso em: 15/12/2006.
- JAMBEIRO, OTHON; BORGES, J.; SANTOS, J. T. J. **Infra-estrutura da Sociedade da Informação: a Indústria de Software em Salvador-BA.** In: *Datagramazero*, v. 7, n. 1, Rio de Janeiro, 2006.
- JAMBEIRO, OTHON; SILVA, H. P. **A Informação e suas Profissões: a Sobrevivência ao Alcance de Todos.** In: *Datagramazero*, v. 5, n. 4, 2004, Rio de Janeiro.
- JEANNE-PERRIER, Valérie; LE CAM, Florence; PÉLISSIER, Nicolas. **Les sites web d'auto-publication : observatoires privilégiés des effervescences et des débordements journalistiques en tous genres.** In: RINGOOT, Roselyne & UTARD, Jean-Michel. *Le journalisme en invention. Nouvelles pratiques, nouveaux acteurs.* Rennes : Presses Universitaires de Rennes, 2005. pp. 161-215.
- JENKINS, Henry; THORBURN, David. **Rethinking Media Change.** The Aesthetics of Transition. Cambridge, Massachusetts: The MIT Press, 2004.
- JOHNSON, Steven. **Cultura da interface.** Como o computador transforma nossa maneira de criar. Rio de Janeiro: Jorge Zahar Editor, 2001.

JORGENSEN, Lee. **Artificial Intelligence: Intelligent Databases**. 2003. In: Applied Knowledge Research Institute. Disponível em: <http://www.akri.org/ai/intdata.htm>. Acesso em: 05/03/2004.

**JORNADAS EM JORNALISMO ON-LINE.2005. ASPECTOS E TENDÊNCIAS**. Universidade da Beira Interior. Covilhã, Portugal, Novembro, 2005.

**JORNADAS DEZ ANOS DE JORNALISMO DIGITAL EM PORTUGAL - ESTADO DA ARTE E CENÁRIOS FUTUROS**. Universidade do Minho. Braga, Portugal, Junho, 2005.

JUAN, Alicia Martinez. **Los retos del periodismo local en la red: hacia una definición del espacio local en la Era global**. In: Sala de Prensa, N° 59, Ano V, Vol. 2. Disponível em: [www.saladeprensa.org](http://www.saladeprensa.org). Acesso em: 19/09/2003.

KODRATOFF, Yves. **About Knowledge Discovery in Texts: A Definition and an Example**. Disponível em: <http://www.cyberlex.pt/y%20kodratoff.pdf>. Acesso em: 16/02/2006.

KAY, Alan; GOLDBERG, Adele. **Personal Dynamic Media**. In: WARDRIP-FRUIIN, Noah; MONTFORT, Nick. **The New Media Reader**. Cambridge-London: MIT Press, 2003. Disponível em: [http://www.mrl.nyu.edu/~noah/nmr/book\\_samples/nmr-26-kay.pdf](http://www.mrl.nyu.edu/~noah/nmr/book_samples/nmr-26-kay.pdf). Acesso em: 13/06/2006.

KISS, Jemina. **Globalnews mapped out**. 2004. Disponível em: [www.journalism.co.uk/news/story857.shtml](http://www.journalism.co.uk/news/story857.shtml). Acesso em: 25/10/2006.

KOCH, Tom. **Journalism for the 21<sup>st</sup> Century**. Online information, electronic databases and the news. New York: Praeger, 1991.

KRAMER, Staci D. **Meet editor Al Gorithm**. 2002. Disponível em: <http://www.ojr.org/ojr/kramer/1033517816.php>. Acesso em: 04/09/2006.

LAGE, Nilson. **JORNALISTAS-ROBÔS**. A era das máquinas inteligentes, ano 1. 2002. Disponível em: <http://observatorio.ultimosegundo.ig.com.br/artigos/da301020021.htm>. Acesso em: 24/08/2006.

\_\_\_\_\_. **A reportagem: teoria e técnica de entrevistas e pesquisa jornalística**. Rio de Janeiro: Record, 2001.

LANCASTER, F.W. **Indexação e resumos: teoria e prática**. Brasília: Briquet de Lemos, 1997. (Título original: *Indexing and abstracting in theory and practice*, 1991).

LANDOW, George P. **Hypertext 2.0 – The Convergence of Contemporary Critical Theory and Technology**. Edição revisada e ampliada do original de 1992. Maryland: The Johns Hopkins University Press, 1997.

- LANINGHAM, Scott. **Tim Berners-Lee**. Originator of the Web and director of the World Wide Web Consortium talks about where we've come, and about the challenges and opportunities ahead. (*Interview*). 28/07/2006. Disponível em: <http://www-128.ibm.com/developerworks/podcast/dwi/cm-int082206.html> e também em: <http://www-128.ibm.com/developerworks/podcast/dwi/cm-int082206.txt> . Acesso em: 26/08/2006.
- LAPHAM, Christine. **The evolution of the newspaper of the future**. Disponível em: <http://sunsite.unc.edu/cm/mag/1995/jul/lapham.html>. Acesso em: 18/07/2001.
- LASICA, J.D. **Balancing Act: How News Portals Serve Up Political Stories**. 2004. Disponível em: <http://ojr.org/ojr/technology/1095977436.php> . Acesso em: 06/05/2006.
- \_\_\_\_\_. **Are digital newspaper editions more than smoke and mirrors?** Disponível em: [www.ojr.org/ojr/technology/1054780452.php](http://www.ojr.org/ojr/technology/1054780452.php) . Acesso em: 25/10/2003.
- LEÃO, Lúcia. **Cartografias em mutação: por uma estética do banco de dados**. In: LEÃO, Lúcia (Org.), **Cibercultura 2.0**. São Paulo: U.N. Nojosa, 2003. pp. 89-105.
- \_\_\_\_\_. **O labirinto da hipermídia**. Arquitetura e navegação no ciberespaço. São Paulo: Iluminuras, 1999.
- LEMOS, André. **Cibercultura**. Alguns pontos para compreender a nossa época. Publicado originalmente em: LEMOS, André; CUNHA, Paulo (Orgs.). **Olhares sobre a Cibercultura**. Porto Alegre: Sulinas, 2003. pp. 11-23. Disponível em: <http://www.facom.ufba.br/ciberpesquisa/andrelemos/cibercultura.pdf>. Acesso em: 17/04/2005.
- \_\_\_\_\_. **Anjos interativos e retribalização do mundo**. Sobre interatividade e interfaces digitais. 1997. Disponível em: <http://www.facom.ufba.br/ciberpesquisa/lemos/interativo.pdf> . Acesso em: 25/10/2004.
- LEVINSON, Paul. **A arma suave**. História natural e futuro da revolução da informação. Lisboa: Bizâncio, 1998.
- LÉVY, Pierre. **Cibercultura**, São Paulo: 34, 1999.
- \_\_\_\_\_. **As tecnologias da inteligência**. O futuro do pensamento na era da informática. São Paulo: 34, 1993.
- LEVY, Mark R. & NEWHAGEN, John E. **The future of Journalism in a Distributed Communication Architecture**. In: BORDEN, Diane L.; HARVEY, Kerric. **The Electronic Grapevine**. Rumor, Reputation, and Reporting in the New On-line Environment. Mahwah, New Jersey, London: LEA, 1998.

- LIEVROUW, Leah A. **What's changed about new media?** Introduction to the fifth anniversary issue of new media & society. In: *New Media & Society*, Vol. 6 (1): 9-15, 2004. Disponível em: [www.sagepublications.com](http://www.sagepublications.com) . Acesso em: 18/03/2004.
- \_\_\_\_\_. **Determination and Contingency in New Media Development: Diffusion of Innovations and Social Shaping of Technology Perspectives.** In: LIEVROUW, Leah A.; LIVINGSTONE, Sonia. **Handbook of new media.** Social Shaping and Consequences of ICTs. London: Sage, 2002. pp. 183-199.
- LIEVROUW, Leah A.; LIVINGSTONE, Sonia. **Handbook of new media.** Social Shaping and Consequences of ICTs. London: Sage, 2002.
- LIMA JÚNIOR, Walter T. **Classificação das bases conceituais para validação de instâncias e atributos a serem utilizados na elaboração de sistema digital de busca de fontes jornalísticas.** 2006. In: Anais do IV SBPJor (CD-ROM). Porto Alegre-RS/Brasil, 2006.
- \_\_\_\_\_. **Pesquisa Aplicada na descoberta de novas tecnologias de produção e distribuição de conteúdos jornalísticos.** 2005. Disponível em: <http://reposcom.portcom.intercom.org.br/bitstream/1904/17377/1/R1800-1.pdf>. Acesso em: 14/05/2006.
- \_\_\_\_\_. **Jornalismo Inteligente (JI) na era do data mining.** In: Anais do II SBPJor (CD-ROM). Salvador-BA/Brasil, 2004.
- LINS DA SILVA, Carlos E.: **Mil dias.** São Paulo: Trajetória cultural, 1988.
- LOCAL WEBLOGS GAWK, STALK – BUT BALK AT MINTING \$\$.** CAN THEY SUCCEED WHERE LOCAL CITY GUIDES FAILED? In: [www.ojr.or/ojr/technology/1054780462.php](http://www.ojr.or/ojr/technology/1054780462.php). Acesso em: 28/10/2003.
- LOPES, Maria Immacolata Vassallo de. **Pesquisa em comunicação.** São Paulo: Loyola, 6ª ed., 2001.
- LÓPEZ, Xosé (Coord.). **Sistemas digitales de información.** Madrid: Pearson Educación, 2006.
- LÓPEZ, Xosé; GAGO, Manuel; PEREIRA, Xosé. **Arquitectura y Organización de la Información.** In: ALIAGA, Ramón Salaverría; DÍAZ NOCI, Javier. **Manual de Redacción Ciberperiodística.** Barcelona: Ariel, 2003.
- LOS GÉNEROS PERIODÍSTICOS EM LOS MEDIOS ONLINE. INFOGRÁFICOS MULTIMEDIA: EL MEJOR EJEMPLO DE NOTICIAS HIPERTEXTUALES** (Maio, 2004). Disponível em: <http://www.mediaccion.com/mediaccionline/temas/periodigital/>. Acesso em: 30/06/2004.

- LLOYD-WILLIAMS, Michael. **Discovering the hidden secrets in your data - the data mining approach to information**. Setembro, 1997. Disponível em: <http://informationr.net/ir/3-2/paper36.html> Acesso em: 09/07/2006.
- LOVINK, Geert. **My first recession**. Critical Internet Culture in Transition. Rotterdam: V2\_Publishing/NAi Publishers, 2003.
- LUBISCO, Nídia; VIEIRA, Sônia C. **Manual de estilo acadêmico**: monografias, dissertações e teses. Salvador: EDUFBA, 2002.
- LUCAS, Luciane. **A hora da 'personalização em massa'**: do consumo de informações à algoritmização do cotidiano. In: *Contemporânea. Revista de Comunicação e Cultura*. V.3, N.2, Abril, 2006. Disponível em: [http://www.contemporanea.poscom.ufba.br/pdfjan2006/contemporanea\\_n3v2\\_lucas.pdf](http://www.contemporanea.poscom.ufba.br/pdfjan2006/contemporanea_n3v2_lucas.pdf). Acesso em: 10/04/2006.
- \_\_\_\_\_. **Comunicação e novas tecnologias**: a mineração de dados e a algoritmização do conhecimento. (Tese de Doutorado). ECO-Pós/UFRJ. Rio de Janeiro, 2004. Disponível em: [http://www.pos.eco.ufrj.br/modules.php?name=Downloads&d\\_op=viewdownload&cid=15](http://www.pos.eco.ufrj.br/modules.php?name=Downloads&d_op=viewdownload&cid=15). Acesso em: 04/08/2006.
- LUKE, Timothy W. **Power and Political Culture**. In: LIEVROUW, Leah A.; LIVINGSTONE, Sonia. **Handbook of new media**. Social Shaping and Consequences of ICTs. London: Sage, 2002.
- LUNDBERG, Jonas. **Layered hypertexts - Adding a new narrative dimension to the web**. Proceedings of The Virtual 2006 - Designing Digital Experience, 14-16. Estocolmo/Suécia, 2006. Também disponível em: [http://m3.sh.se/get\\_file.php?id\\_file=62](http://m3.sh.se/get_file.php?id_file=62). Acesso em: 14/12/2006.
- LYOTARD, Jean-François. **A condição pós-moderna**. Rio de Janeiro: José Olympio, 1998.
- MACHADO, Elias. **O Jornalismo Digital em Base de Dados**. Florianópolis: Calandra, 2006.
- \_\_\_\_\_. **O banco de dados como espaço de composição de narrativas multimídia**. In: Anais do II SBPJor. Salvador-BA/Brasil, 2004b.
- \_\_\_\_\_. **Banco de dados como formato no jornalismo digital**. In: Anais III Sopcom, VI Lusocom, II Ibérico, UBI (CD-ROM), 2004a.
- \_\_\_\_\_. **A Intranet como Modelo de Gestão no Jornalismo Digital**. In: *Pauta Geral Revista de Jornalismo*. Ano 11 – Nº 6. Salvador: Calandra, 2004c.
- \_\_\_\_\_. **O ciberespaço como fonte para os jornalistas**. Salvador: Calandra, 2003 (Coleção Biblioteca J).

- \_\_\_\_\_. **O jornal digital como epicentro das redes de circulação de notícias.** In: *Pauta Geral*, Nº4, Salvador: Calandra, 2002.
- \_\_\_\_\_. **La Estructura de la Noticia en las Redes Digitales:** un estudio de las consecuencias de las metamorfosis tecnológicas en el periodismo. Tese Doutoral em Comunicação – Universidad Autónoma de Barcelona, Barcelona, 2000.
- MACHADO, Elias; PALACIOS, Marcos. **Um modelo híbrido de pesquisa:** a metodologia aplicada pelo GJOL. Artigo inédito a ser publicado In: **Metodologias de Pesquisa em Jornalismo.** In: MACHADO, Márcia Benetti Machado; LAGO, Cláudia. (Orgs.). No prelo, 2005.
- \_\_\_\_\_. **Modelos de Jornalismo Digital.** Salvador: Calandra, 2003.
- MAINGUENEAU, Dominique. **Análise de textos de comunicação.** São Paulo: Cortez, 2002.
- MANOVICH, Lev. **Alan Kay's Universal Media Machine.** 2006. Disponível em: [www.manovich.net/DOCS/kay\\_article.doc](http://www.manovich.net/DOCS/kay_article.doc). Acesso em 03/10/2006.
- \_\_\_\_\_. **Social data browsing.** Fev/2006. Disponível em: <http://www.tate.org.uk/netart/bvs/manovich.htm>. Acesso em: 03/10/2006.
- \_\_\_\_\_. **Image Future.** In : *Animation interdisciplinary journal*. Vol. 1, Nº. 1, 25-44 (2006). Disponível em: <http://anm.sagepub.com/cgi/reprint/1/1/25>. Acesso em: 12/7/2006.
- \_\_\_\_\_. **Understanding Meta-Media.** *Ctheory.net*. 2005. Disponível em: <http://www.ctheory.net/articles.aspx?id=493>. Acesso em: 17/07/2006.
- \_\_\_\_\_. **Quem é o autor?** Sampleamento/Remixagem/Código aberto. In: BRASIL, André (et al) *Cultura em fluxo : novas mediações em rede.* Belo Horizonte: PUC Minas, 2004. pp: 248-263
- \_\_\_\_\_. **The language of new media.** Cambridge: MIT Press, 2001.
- \_\_\_\_\_. **New Media from Borges to HTML.** In: WARDROP-FRUIIN, Noah; MONTFORT, Nick. **The New Media Reader.** Cambridge-London: MIT Press, 2003. Disponível em: [http://www.mrl.nyu.edu/~noah/nmr/book\\_samples/nmr-intro-manovich-excerpt.pdf](http://www.mrl.nyu.edu/~noah/nmr/book_samples/nmr-intro-manovich-excerpt.pdf). Acesso em: 13/06/2006.
- \_\_\_\_\_. **Database as a genre of new media.** In: *AI Soc.* 14, 2 (May. 2000), 176-183. Disponível em: <http://dx.doi.org/10.1007/BF01205448>). Acesso em: 27/07/2006. Também disponível em: [http://vv.arts.ucla.edu/publications/publications/98-99/ai\\_society/manovich.html](http://vv.arts.ucla.edu/publications/publications/98-99/ai_society/manovich.html). 1999. Acesso em: 27/07/2006.
- MANOVICH, Lev; KRATKY, Andreas. **Soft Cinema.** Navigating the database. DVD-video with 40 page color booklet. Cambridge: The MIT Press, 2005.

- MARCOS, José Alvarez. **El periodismo ante la tecnología hipertextual**. In: DIAZ NOCI, Javier; SALAVERRÍA, Ramón. **Manual de Redación Ciberperiodística**. Barcelona: Ariel, 2003
- MARCOS RECIO, Juan Carlos. **La Documentación Electrónica en los Medios de Comunicación**. Madrid: Editorial Fragua, 1999.
- MARSHALL, P. David. **New Media Cultures (Cultural Studies in Practice)**. Hodder: Arnold Publication, 2004.
- MAURER, Herman; SCHER BACOV, Nick.; HALIM, Zahram; RAZAK, Zaidah. **From Databases to Hypermedia: with 26 CAI Lessons**. Berlin: Springer, 1998.
- MAZZIOTTI, Nora. **Narrativa: gêneros na televisão pública**. In: RINCÓN, Omar (org.). **Televisão pública: do consumidor ao cidadão**. São Paulo: Friedrich-Ebert-Stiftung, 2002.
- MCADAMS, Mindy. **Flash Journalism**. How to create multimedia news packages. Burlington: Focal Press, 2005.
- \_\_\_\_\_. **Inventing Online newspaper**. Disponível em: [www.sentex.net/~mmcadams/invent.html](http://www.sentex.net/~mmcadams/invent.html), (publicado pela primeira vez em 1995 no Interpersonal Computing and Technology: as electronic journal for the 21st century). ISSN: 1064-4326, July 1995, v.3, pp.64-90. Acesso em: 23/09/2002.
- MCGARRY, Kevin. **Database Imaginary: Sarah Cook, Steve Dietz, and Anthony Kiendl interviewed by Kevin McGarry**. Disponível em: <http://rhizome.org/thread.rhiz?thread=15882&page=1#30434>. Acesso em: 25/09/2006.
- MCLUHAN, Marshall. **Os meios de comunicação como extensões do homem**. (Understanding Media: The Extensions of Man). Trad. Décio Pignatari. São Paulo, Cultrix, 1969.
- MCLUHAN, Marshal; FIORE, Quentin. **The medium is the message**. An inventory of effects. Harmondsworth, Allen Lane: The Penguin, Press, 1967.
- MEADOWS, Mark Stephen. **Pause & Effect**. The art of interactive narrative. Indiana: New Riders, 2003.
- MELLER, Paul. **Corte obriga Google a retirar do serviço News jornais da Bélgica**. *IDG Now*, 18/09/2006. Disponível em: [http://pcworld.uol.com.br/noticias/2006/09/19/idgnoticia.20060919.7019011743/IDGNoticia\\_view](http://pcworld.uol.com.br/noticias/2006/09/19/idgnoticia.20060919.7019011743/IDGNoticia_view). Acesso em: 30/10/2006.
- MELO, José Marques de. **A opinião no jornalismo brasileiro**. Petrópolis: Vozes, 1985.

MEYER, Phillip. **The future of CAR: Declare Victory and Get Out!** From “When Nerds and World Collide: Reflections on the Development of Computer Assisted Reporting. 1999. *The Poynter Institute for Media Studies*. Disponível em: [www.unc.edu/~pmeyer/carfuture.doc](http://www.unc.edu/~pmeyer/carfuture.doc). Acesso em: 17/05/2006.

\_\_\_\_\_. **Periodismo de Precisión: Nuevas Fronteras para La Investigación Periodística**. Barcelona: Bosch, 1993.

\_\_\_\_\_. **The new precision journalism**. Bloomington: Indiana University Press, 1991.

MIELNICZUK, Luciana; MARQUES, Iuri Lamel. **Sistemas publicadores para webjornalismo: MapaLink, um protótipo para produtos de terceira geração**. In: Anais XV Encontro Anual da Compós (CD-ROM). Bauru: Editora UNESP, 2006, V. 1.

MIELNICZUK, Luciana. **Jornalismo na Web: uma contribuição para o estudo do formato da notícia na escrita hipertextual**. (Tese de Doutorado). FACOM/UFBA, Salvador, 2003.

MIELNICZUK, Luciana; PALACIOS, Marcos. **Considerações para um estudo sobre o formato da notícia na Web: o link como elemento paratextual**. In: Anais X Encontro Nacional da COMPÓS (GT Jornalismo). Brasília, Junho, 2001.

MIRANDA, Eduardo Ferreira de. **PIP Panopticon: Memória do Desenvolvimento do Sistema de Publicação para o Panopticon**. (Trabalho de TCC). FACOM/UFBA, 2006.

MOHERDAUI, Luciana. **O Usuário de Notícias no Jornalismo Digital: um estudo sobre a função do sujeito no Último Segundo e no A Tarde Online**. (Dissertação de Mestrado). FACOM/UFBA, Salvador, 2005.

\_\_\_\_\_. **Guia de estilo Web - Produção e edição de notícias online**. São Paulo: Senac, 2000.

MONMONIER, Mark. **Maps in the electronic media**. In: **Maps with the news**. The development of American journalistic cartography. Chicago: University of Chicago Press, 1989.

MORGAINÉ, Daniel. **Diez años para sobrevivir: el diario de masas de 1980**. Madrid: Editora Nacional, 1971.

MORRIS, Merrill; OGAN, Christine. **The Internet as Mass Medium**, Disponível em: <http://207.201.161.120/jcmc/vol1/issue4/>. Acesso em: 09/07/97.

MOUILLAUD, Maurice; PORTO, Sérgio Dayrell (org.). **O jornal: da forma ao sentido**. Brasília: Paralelo 15, 1997.

- MOURA, Catarina. **O jornalismo na era Slashdot**. In: FIDALGO, António; SERRA, Paulo (Orgs.). **Informação e Comunicação Online**. Jornalismo Online. Volume 1. Covilhã: Universidade da Beira Interior/Portugal, 2003. pp: 189-198.
- MOURIQUAND, Jacques. **O jornalismo de investigação**. Coleção Técnicas de Jornalismo. Lisboa: Editorial Inquérito, 2002. Título original: *L'enquête*. Tradução: Mário Matos e Lemos.
- MURRAY, Janet H. **Inventing the Medium**. In: WARDRIP-FRUIN, Noah; MONTFORT, Nick. **The New Media Reader**. Cambridge-London: MIT Press, 2003b. Disponível em: [http://www.mrl.nyu.edu/~noah/nmr/book\\_samples/nmr-intro-murray-excerpt.pdf](http://www.mrl.nyu.edu/~noah/nmr/book_samples/nmr-intro-murray-excerpt.pdf). Acesso em: 13/06/2006.
- \_\_\_\_\_. **Hamlet no Holodeck**. O futuro da narrativa no ciberespaço. São Paulo: Itaú Cultural: Unesp, 2003a.
- \_\_\_\_\_. **Hamlet on the Holodeck**. The Future of Narrative in Cyberspace. New York: The MIT Press, 1997.
- \_\_\_\_\_. **Hamlet en la holocubierta**. El futuro de la narrativa en el ciberespacio. Barcelona: Paidós, 1999.
- NAVEGA, Sérgio. **Princípios essenciais do data mining**. Anais do Infoimagem. Novembro, 2002. Disponível em: [www.intelliwise.com/reports/i2002.htm](http://www.intelliwise.com/reports/i2002.htm). Acesso em: 04/08/2006.
- NEGROPONTE, Nicholas. **Vida Digital**. São Paulo: Companhia das Letras, 1995.
- NEWHAGEN, John E.; LEVY, Mark R. **The Future of Journalism in a Distributed Communication Architecture**. In: D. L. Borden and H. Kerric (Eds.), *The Electronic Grapevine: Rumor, reputation and reporting in the new on-line environment*, Mahwah: Lawrence Erlbaum, 1998. pp. 9-21.
- NIELSEN, Jacob. **Projetando websites**. Rio de Janeiro: Campus, 2000. (Título original: *Designing web usability*).
- NILES, Robert. **“The programmer as journalist: a Q&A with Adrian Holovaty”**. In: OJR, Online Journalism Review. Junho, 2006. Disponível em: <http://www.ojr.org/ojr/stories/060605niles/>. Acesso em: 30/06/2006.
- NOGUEIRA, Leila. **O webjornalismo audiovisual: uma análise de notícias no UOL News e na TV UERJ Online**. (Dissertação de Mestrado). FACOM/UFBA, Salvador 2005.
- NOGUEIRA, Luís. **Jornalismo na rede: arquivo, acesso, tempo, estatística e memória**. In: FIDALGO, António; SERRA, Paulo (Orgs.). **Informação e Comunicação Online**. Jornalismo Online. Volume 1. Covilhã: Universidade da Beira Interior/Portugal, 2003. pp: 159-178.

- \_\_\_\_\_. **O mundo como base de dados.** 2001. Disponível em: [www.bocc.ubi.pt](http://www.bocc.ubi.pt). Acesso em: 18/03/2005.
- ORDUÑA, Octavio I. R. et al. **Blogs - La conversación en Internet que está revolucionando medios, empresas y a ciudadanos.** Madrid: ESIC, 2005.
- O'REILLY, Tim. **What is Web 2.0.** Design Patterns and Business Models for the Next Generation of Software. Disponível em: <http://www.oreillynet.com/pub/a/oreilly/tim/news/2005/09/30/what-is-web-20.html>. Acesso em: 03/12/2006.
- ORIGGI, Gloria; ARIKHA, Noga. **Le texte à l'heure de l'internet.** Paris : Bibliothèque publique d'information, 2003.
- ORIHUELA, José Luís. **Weblogs y blogosfera: el medio y la comunidad.** In: ORDUÑA, Octavio I. R. et al. **Blogs - La conversación en Internet que está revolucionando medios, empresas y a ciudadanos.** Madrid: ESIC, 2005. pp: 13-48.
- ORTUÑO, Pedro Manuel Díaz. **Problemática y Tendencias en la arquitectura de metadatos web.** Disponível em: [www.um.es/fccd/anales/ado6/ad0603.pdf](http://www.um.es/fccd/anales/ado6/ad0603.pdf). Acesso em: 26/07/2006.
- OUTING, Steve. **Why Aren't Newspapers Breaking Out of the Box?** In: Editor & Publisher, Setembro-2006. Disponível: [http://www.editorandpublisher.com/eandp/columns/stopthepresses\\_display.jsp?vnu\\_content\\_id=1003156892](http://www.editorandpublisher.com/eandp/columns/stopthepresses_display.jsp?vnu_content_id=1003156892). Acesso em: 29/09/2006.
- \_\_\_\_\_. **How to Make Your Web Site More Conversational.** In: Editor & Publisher, Junho-2006. Disponível em: [http://www.editorandpublisher.com/eandp/columns/stopthepresses\\_display.jsp?vnu\\_content\\_id=1002728381](http://www.editorandpublisher.com/eandp/columns/stopthepresses_display.jsp?vnu_content_id=1002728381). Acesso em: 23/07/2006.
- \_\_\_\_\_. **A decade of progress in online news but we're still making some of the same mistakes.** In: Editor & Publisher, dezembro-2003. [mensagem pessoal]. Mensagem recebida por [suzana-b@uol.com.br](mailto:suzana-b@uol.com.br) em 25/06/2004.
- PALACIOS, Marcos. **Natura non facit saltum: Promessas, alcances e limites no desenvolvimento do jornalismo on-line e da hiperficcão.** *e-Compós.* Revista da Associação Nacional dos Programas de Pós-Graduação em Comunicação. Edição 2-Abril, 2005. Disponível em: [http://www.compos.org.br/e-compos/adm/documentos/abril2005\\_palacios.pdf](http://www.compos.org.br/e-compos/adm/documentos/abril2005_palacios.pdf). Acesso em 15/05/2005.
- \_\_\_\_\_. **Mundo Digital.** In: RUBIM, Albino (Org.). **Cultura e Atualidade no Vestibular.** Salvador: EDFUBA, 2005. Também disponível em: <http://www.facom.ufba.br/jol/producao.htm>. Acesso em: 03/08/2006.
- \_\_\_\_\_. **Jornalismo em ambiente plural?** Notas para discussão da internet enquanto suporte para a prática jornalística. In: BRASIL, André (et al). **Cultura em fluxo: novas mediações em rede.** Belo Horizonte: PUC Minas, 2004. pp: 248-263.

- \_\_\_\_\_. **Internet as system and environment in cyberspace:** Preliminary ideas from ongoing research. In: *Triple C* 1 (2): 95-104, 2003c. Disponível em: <http://triplec.uti.at/articles>. Acesso em: 23/11/2003.
- \_\_\_\_\_. **Ruptura, continuidade e potencialização no jornalismo on-line:** o lugar da memória. In: MACHADO, Elias e PALACIOS, Marcos. **Modelos de Jornalismo Digital**. Salvador: Calandra, 2003b.
- \_\_\_\_\_. **Jornalismo online, informação e memória:** apontamentos para debate. Apresentado nas Jornadas de Jornalismo Online, Departamento de Comunicação e Artes, Universidade da Beira Interior (Portugal), Junho, 2002. In: FIDALGO, António; SERRA, Paulo (Orgs.). **Informação e Comunicação Online**. Jornalismo Online. Volume 1. Covilhã: Universidade da Beira Interior/Portugal, 2003a.
- \_\_\_\_\_. **Jornalismo online, informação e memória:** apontamentos para debate. Apresentado nas Jornadas de Jornalismo Online, Departamento de Comunicação e Artes, Universidade da Beira Interior (Portugal), Junho, 2002a. Disponível em: [www.facom.ufba.br/jol/producao.htm](http://www.facom.ufba.br/jol/producao.htm). Acesso em: 17/08/2003.
- \_\_\_\_\_. **Fazendo jornalismo em redes híbridas.** Notas para discussão da internet enquanto suporte mediático. Texto publicado no Observatório da Imprensa em 11/12/2002. 2002b. Disponível em: [www.observatoriodaimprensa.com.br/artigos/eno111220022.htm](http://www.observatoriodaimprensa.com.br/artigos/eno111220022.htm). Acesso em: 03/08/2003.
- \_\_\_\_\_. **Hipertexto, fechamento e o uso do conceito de não-linearidade discursiva.** 1999. Disponível em: <http://www.facom.ufba.br/ciberpesquisa/palacios/hipertexto.html>. Acesso em: 25/10/2000.
- \_\_\_\_\_. **O que há de (realmente) novo no Jornalismo Online?** Conferência proferida por ocasião do concurso público para Professor Titular na FACOM/UFBA, 21.09.1999.
- PALACIOS, Marcos; MIELNICZUK, Luciana; BARBOSA, Suzana; RIBAS, Beatriz; NARITA, Sandra. **Um mapeamento de características e tendências no jornalismo online brasileiro e português.** In: *Comunicarte, Revista de Comunicação e Arte*, Vol.1, N°2, Universidade de Aveiro, Portugal, Setembro/2002.
- PALACIOS, Marcos; MACHADO, Elias. **Manual do jornalismo na internet.** Noções básicas, conceitos e um guia sobre as principais publicações jornalísticas digitais no Brasil e no Exterior. Salvador: FACOM-UFBA, 1997. Disponível em [http://www.facom.ufba.br/jol/fontes\\_manuais.htm](http://www.facom.ufba.br/jol/fontes_manuais.htm). Acesso em: 20/11/2000.
- PAUL, Nora. **'New News' retrospective:** Is online news reaching its potential? Disponível em: <http://www.ojr.org/ojr/stories/050324paul>. Acesso em: 20/04/2005.

- \_\_\_\_\_. **Computer-Assisted Research.** A Guide to Tapping Online Information. Chicago: Bonus, 1999. (Fourth Edition).
- PAUL, Nora; FIEBICH, Christina. **Five Elements of digital storytelling.** Disponível em: [www.inms.umn.edu/Elements/index.htm](http://www.inms.umn.edu/Elements/index.htm). Acesso em: 04/06/2002.
- PAUL, Christiane: "**The Database as System and Cultural Form: Anatomies of Cultural Narratives**". In: Database Aesthetics, Victoria Vesna (Hg.), 2004 (forthcoming). Disponível em: [www.cityarts.com/paulc/RISD/Paul\\_Database.doc](http://www.cityarts.com/paulc/RISD/Paul_Database.doc). Acesso em: 27/07/2006.
- PARK, Robert. **News and opinion.** Collected Papers III. Society. Glencoe: Free Press, 1955.
- PAVLIK, John V. **El periodismo y los nuevos medios de comunicación.** Barcelona: Paidós Comunicación, 2005.
- \_\_\_\_\_. **Journalism and new media.** New York: Columbia University Press, 2001.
- PEREIRA, Xosé. **La presencia gráfica del sistema (front-end).** In: LÓPEZ, Xosé. **Sistemas digitales de información.** Madrid: Pearson Educación, 2006. pp: 143-215.
- PEREIRA, José Luís. **Tecnologia de bases de dados.** Lisboa: FCA Editora de Informática, 1998 (2ª ed.).
- PESQUISA SOBRE O USO DAS TECNOLOGIAS DA INFORMAÇÃO E DA COMUNICAÇÃO NO BRASIL 2005.** Comité Gestor da Internet no Brasil/CGI. Disponível em: <http://www.nic.br/indicadores/indicadores.pdf>. Acesso em: 23/08/2006.
- PIATTINI, Mario. **Líneas de evolución de las bases de datos.** In: NOVATICA / may.-jun. 2000 / Especial 25 aniversario. Disponível em: [www.ati.es/novatica/2000/145/marpia-145.pdf](http://www.ati.es/novatica/2000/145/marpia-145.pdf). Acesso em 29/09/2006.
- PIATTINI, Mário; CALERO, Coral. **Bases de Datos Activas.** 1999. Disponível em: <http://alarcos-cr.uclm.es/doc/bbddavanzadas/activas.pdf>. Acesso em: 29/09/2006.
- PINTO, Álvaro Vieira. **O conceito de tecnologia.** Volume 1. Rio de Janeiro: Contraponto, 2005.
- PISCITELLI, Alejandro. **Internet, la imprenta del siglo XXI.** Barcelona: Gedisa, 2005.
- PORT DA ROCHA, Rafael. **Metadados, Web Semântica, Categorização Automática:** combinando esforços humanos e computacionais para a descoberta e uso dos recursos da web. In: *Em Questão*. Porto Alegre, V.10, Nº. 1, Jan./Jun., 2004. Disponível em:

[http://www6.ufrgs.br/emquestao/pdf\\_2004\\_v10\\_n1/EmQuestaoV10\\_N1\\_2004\\_abertura.pdf](http://www6.ufrgs.br/emquestao/pdf_2004_v10_n1/EmQuestaoV10_N1_2004_abertura.pdf). Acesso em: 17/07/2006.

\_\_\_\_\_ et al. **Base de Dados como Elemento Facilitador para o Tratamento dos Dados Terminológicos.** In: *Em Questão*, Porto Alegre, V.10, Nº 2, Jul/Dez., 2004. Disponível em: [http://www6.ufrgs.br/emquestao/pdf\\_2004\\_v10\\_n2/EmQuestaoV10\\_N2\\_2004\\_rel02.pdf](http://www6.ufrgs.br/emquestao/pdf_2004_v10_n2/EmQuestaoV10_N2_2004_rel02.pdf). Acesso em: 17/07/2006.

POSTER, Mark. **The second media age.** Cambridge: Polity Press, 1995.

\_\_\_\_\_. **A segunda era dos media.** Oeiras: Celta Editora, 2000.

PRADO, Ana Lucia. **Uma notícia a cada 90 segundos:** estudo de caso de um jornal *online*. Pauta Geral, ano 9, n.4, Ed. Calandra, 2002.

PRYOR, Larry. **It feels relevant:** biological tactility in news media. *Online Journalism Review*, 24/setembro/2006. Disponível em: <http://www.ojr.org/ojr/stories/060925pryor/>. Acesso em: 27/09/2006.

\_\_\_\_\_. **The third wave of online journalism.** *Online Journalism Review*, 18/abril/2002. Disponível em: [www.ojr.org/ojr/future/1019174689.php](http://www.ojr.org/ojr/future/1019174689.php). Acesso em 28/10/2003.

PUJOL, Maria Eulália Fuentes I. (editora). **Manual de documentación periodística.** Madrid: Editorial Síntesis, 1995.

\_\_\_\_\_. **Acceso y recuperación de información en bases de datos en línea/Internet.** Disponível em: <http://www.eubd.ucm.es/html/personales/enred/mantonia/docauto/tema14/tema14.html>. Acesso em 10/03/2005.

QUADROS, Claudia Irene de. **Base de dados:** a memória extensiva do jornalismo. In: *Em Questão*. Porto Alegre. V. 11. N.2, Jul./dez., 2005.

\_\_\_\_\_. **Jornalismo e base de dados para gerar conhecimento.** In: Anais do II SBPJor. Salvador-BA/Brasil, 2004.

QUADROS, Claudia Irene de; ROSA, Ana Paula da; VIEIRA, Josiany. **Blogs e as transformações do Jornalismo.** *e-Compós*. Revista da Associação Nacional dos Programas de Pós-Graduação em Comunicação. Edição 3-Agosto, 2005. Disponível em: [http://www.assimcomunicacao.com.br/revista/documentos/agosto2005\\_quadrosrosvieira.pdf](http://www.assimcomunicacao.com.br/revista/documentos/agosto2005_quadrosrosvieira.pdf). Acesso em 24/10/2005.

QUIJADA, David Fernandez. **Bases de datos multimédia en la red global.** Un recurso para la localización e indexación de contenidos. In: Telos: Cuadernos de comunicación, tecnología, y socialidad. Abril/Junho-2006, Nº67, págs. 28-33. Disponível em:

[www.campusred.net/telos/articuloperspectiva.asp?idarticulo=18rev=67](http://www.campusred.net/telos/articuloperspectiva.asp?idarticulo=18rev=67). Acesso em: 26/07/2006.

QUINN, Stephen. **Convergent journalism**. The fundamentals of Multimedia Reporting. New York: Peter Lang Publishing, 2005.

\_\_\_\_\_. **Knowledge management in the digital newsroom**. London: Focal Press, 2002.

REAVY, Matthew M. **Introduction to Computer-Assisted Reporting**. A Journalist's Guide. California: Mayfield Publishing Company, 2001.

RIBAS, Beatriz. **A Narrativa Webjornalística** - um estudo sobre modelos de composição no ciberespaço. (Dissertação de Mestrado). FACOM/UFBA, Salvador 2005. Disponível em: [www.facom.ufba.br/jol/producao\\_dissertacoes](http://www.facom.ufba.br/jol/producao_dissertacoes). Acesso em: 27/06/2005.

RINGOOT, Roselyne & UTARD, Jean-Michel. **Le journalisme en invention**. *Nouvelles pratiques, nouveaux acteurs*. Rennes : Presses Universitaires de Rennes, 2005.

RIVERO, Raúl. **Elmundo.es y su planteamiento técnico**. Disponível em: [www.elmundo.es/imasd/docs/charlas/2006-caceres/elmundo-es-y-su-planteamiento-tecnico.pdf](http://www.elmundo.es/imasd/docs/charlas/2006-caceres/elmundo-es-y-su-planteamiento-tecnico.pdf). Acesso em: 02/11/2006.

RODRIGUES, Adriano Duarte. **As novas tecnologias da informação e a experiência**. Disponível em : [www.bocc.ubi.pt/pag/rodrigues-adriano-comunicacao-experiencia.htm](http://www.bocc.ubi.pt/pag/rodrigues-adriano-comunicacao-experiencia.htm). Acesso em 13/07/2006.

ROGERS, Everett M. **Diffusion of innovations**. (4<sup>th</sup> edition). New York: Free Press, 1995.

ROSEN, Jay. **The era of networked journalism begins**. In: *PressThink*. August, 2006a. Disponível em: [http://journalism.nyu.edu/pubzone/weblogs/pressthink/2006/08/15/ear\\_ntw.html](http://journalism.nyu.edu/pubzone/weblogs/pressthink/2006/08/15/ear_ntw.html). Acesso em: 27/09/2006.

\_\_\_\_\_. **Web users open the gates**. June, 2006b. In: *washingtonpost.com*. Disponível em: <http://www.washingtonpost.com/wpdyn/content/article/2006/06/18/AR2006061800618.html> . Acesso em: 21/09/2006..

\_\_\_\_\_. **Bloggers vs Journalists is Over**. In: *PressThink*. January, 2005a. Disponível em: [http://journalism.nyu.edu/pubzone/weblogs/pressthink/2005/01/15/berk\\_pprd.html](http://journalism.nyu.edu/pubzone/weblogs/pressthink/2005/01/15/berk_pprd.html) l. Acesso em: 27/07/2006.

\_\_\_\_\_. **Laying the Newspaper Gently Down to Die**. In: *PressThink*. March, 2005b. Disponível em: 27/07/2006.

- \_\_\_\_\_. **Emerging alternatives. Terms of authority.** In: *Columbia Journalism Review*. 2003. Disponível em: <http://www.cjr.org/issues/2003/5/alt-rosen.asp> . Acesso em 21/09/2006.
- RYAN, Marie-Laure. **Narrative as virtual reality.** Immersion and interactivity in literature and electronic media. Baltimore: The Johns Hopkins University Press, 2001.
- SÁ, Alberto. **O código oculto nas notícias eletrônicas.** In: Actas 4º SOPCOM. Aveiro, Portugal, Outubro, 2005.
- SALAVERRÍA, Ramón. (Coord.). [Cibermedios. El impacto de internet en los medios de comunicación en España.](#) Sevilla: Comunicación Social Ediciones y Publicaciones, 2005b.
- \_\_\_\_\_. **Redación periodística en internet.** Barcelona: EUNSA, 2005a.
- \_\_\_\_\_. **O jornalismo digital num mundo em constante mudança.** Conferência proferida nas jornadas “Dez Anos de Jornalismo Digital em Portugal – Estado da Arte e Cenários Futuros” ([www.dezanos.blogspot.com](http://www.dezanos.blogspot.com)), organizadas pelo projecto Mediascópio, do Centro de Estudos de Comunicação e Sociedade da Universidade do Minho (Portugal), 03 de Junho, 2005c.
- SALAVERRÍA, Ramon; CORES, Rafael. **Géneros periodísticos en los cibermedios hispanos.** In: SALAVERRÍA, Ramón. (Coord.). [Cibermedios. El impacto de internet en los medios de comunicación en España.](#) Sevilla: Comunicación Social Ediciones y Publicaciones, 2005a. p: 145-185.
- SALGADO, Silvia Eloisa Carlín; RODRIGUEZ, Rosendo Moreno. **Valorización de Las Bases de Datos Deductivas.** In: *Quaderns Digitals / Quaderns 31* . 07-12-2003. Disponível em: [http://www.quadernsdigitals.net/index.php?accionMenu=hemeroteca.VisualizaArticuloIU.visualiza&articulo\\_id=6850](http://www.quadernsdigitals.net/index.php?accionMenu=hemeroteca.VisualizaArticuloIU.visualiza&articulo_id=6850). Acesso em: 29/09/2006.
- SANTAELLA, Lucia. **Comunicação e pesquisa.** Projetos para mestrado e doutorado. São Paulo: Hacker Editores, 2001.
- SANCHO, José Luís Valero. **El relato en la infografía digital.** In: DÍAZ NOCI; SALAVERRÍA. **Manual de Redación Ciberperiodística.** Barcelona: Ariel, 2003.
- SANTOS, Milton. **A natureza do espaço: técnica e tempo - razão e emoção.** São Paulo: Hucitec, 1996.
- SARKAR, Dibya. **Going where no search engine has gone before.** Connotate Technologies uses information agents to extract data from Deep Web. In: FCW.Com, Maio, 2005. Disponível em: <http://www.fcw.com/article88982-05-30-05-Print>. Acesso em: 28/11/2005.

- SCHEDDEN, David. **New Media Timeline (1969-1998)**. *Poynteronline*. 1998. Disponível em: [www.poynter.org](http://www.poynter.org). Acesso em: 01/08/2004.
- SCHWINGEL, Carla; MACHADO, Elias; PALACIOS, Marcos; ROCHA, Lucas. **Um Jornal Laboratório multimídia, multi-usuário e descentralizado**. O caso da Plataforma Panopticon. In: *Pauta Geral*, Ano 12, Nº 7. Salvador: Calandra, 2005. pp: 105-115.
- SCHWINGEL, Carla. **Jornalismo digital de quarta geração: a emergência de sistemas automatizados para o processo de produção industrial no jornalismo digital**. In: Anais do XIV Encontro Anual da Compós, 2005. Também disponível em: [www.ufrgs.br/gtjornalismocompos/estudos2005.htm](http://www.ufrgs.br/gtjornalismocompos/estudos2005.htm). Acesso em: 15/02/2006.
- \_\_\_\_\_. **Os sistemas de publicação como fator da terceira geração do jornalismo digital**. 2004. Disponível em: [www.facom.ufba.br/jol/pdf/2004\\_schwingel\\_sistemas\\_publicacao.pdf](http://www.facom.ufba.br/jol/pdf/2004_schwingel_sistemas_publicacao.pdf). Acesso em: 16/05/2005.
- \_\_\_\_\_. **A teoria na prática no Jornal Experimental Panopticon**. In: MACHADO, Elias e PALACIOS, Marcos. **Modelos de Jornalismo Digital**. Salvador: Calandra, 2003.
- SCHULTZ, Tanjev. **Interactive Options in Online Journalism**. Disponível em: <http://www.ascusc.org/jcmc/vol5/issue1/schultz.html>. Acesso em: 14/09/2000.
- SCOLARI, Carlos. **Hacer Clic**. Hacia una sociosemiótica de las interacciones digitales. Barcelona: Gedisa Editorial, 2004.
- SCOLARI, Carlos; MARCH, Jaume. **Hacia una taxonomía de los regímenes de info-visualización**. In: *Interacción*. Maio, 2004, Lleida/Espanha. pp: 01-07. Disponível também em: <http://griho.udl.es/i2004/BajarPonencia/9.pdf>. Acesso em: 10/08/2005.
- SEIXAS, Lia. **A rosa dos gêneros**. Uma proposição de análise dos elementos cardeais dos gêneros jornalísticos. (Exame de Qualificação). FCOM/UFBA, Salvador, 2006.
- \_\_\_\_\_. **Os gêneros jornalísticos digitais**. Um estudo das práticas discursivas no ambiente digital. In: Anais do XIII Encontro Anual da Compós, 2004. Também disponível em: <http://www.ufrgs.br/gtjornalismocompos/estudos2004.htm>. Acesso em: 18/06/2005.
- SELLARÉS, M<sup>a</sup>. José Recoder. **El videotext y la prensa**. In: PUJOL, Maria Eulália Fuentes I. (editora). **Manual de documentación periodística**. Madrid: Editorial Síntesis, 1995. pp: 180-194.
- SEMONCHE, Barbara P. **History of News Libraries**. (Excerpted from SEMONCHE, Barbara P. (ed.). *News Media Libraries: A Management Handbook*. Greenwood Press, 1993a. Disponível em: <http://parklibrary.jomc.unc.edu/newslibhist2.html> . Acesso em 02/08/2006.

- \_\_\_\_\_. **Computer Assisted Journalism**. (Excerpted from SEMONCHE, Barbara P. (ed.). *News Media Libraries: A Management Handbook*. Greenwood Press, 1993b. Disponível em: <http://parklibrary.jomc.unc.edu/caj.html> . Acesso em 02/08/2006.
- SHNEIDERMAN, Ben. **Foreword: Information Visualization**. Beyond the horizon (Chen, 2004). Disponível em: <http://www.cs.umd.edu/users/ben/Chen-InfoViz-book-foreword.pdf>. Acesso em: 04/11/2006.
- SHNEIDERMAN, Ben; JOHNSON, Brian. **Tree-Maps: A Space-Filling Approach to the Visualization of Hierarchical Information Structures**. Disponível em: <http://delivery.acm.org/10.1145/950000/949654/p284-johnson.pdf?key1=949654&key2=0330199611&coll=&dl=ACM&CFID=15151515&CFTOKEN=6184618>. Acesso em: 04/11/2006.
- SILVA, Mário J. **ARIADNE**. Relatório final de execução. 2000. Disponível em: <http://xldb.fc.ul.pt/ariadne/documentos/ariadne-relatorio-execucao-final.pdf>. Acesso em 08/09/2006.
- SILVA JR., José Afonso da. **Uma trajetória em redes: modelos e características operacionais das agências de notícias, das origens às redes digitais, com três estudos de caso**. (Tese de Doutorado). FACOM/UFBA, Salvador, 2006.
- \_\_\_\_\_. **Jornalismo 1.2: características e usos da hipermídia no jornalismo, com estudo de caso do Grupo Estado de São Paulo**. (Dissertação de Mestrado). FACOM/UFBA, Salvador, 2000.
- SIMÃO, João. **Akademia, um modelo de webjornal**. 2005. Disponível em: <http://www.bocc.ubi.pt/pag/simao-joao-akademia-modelo-webjornal.pdf>. e também em: <http://simaocc.home.sapo.pt/>. Acesso em: 27/05/2005.
- SIMONE, José Fernando; MONTEIRO, Mariana. **Jornalismo online - O futuro da informação**. Rio de Janeiro: Webmeio Edições, 2001.
- SIQUEIRA, Ethevaldo. **Informação eletrônica e novas tecnologias no Brasil**. (capítulo especial). In: CODINA, Lluís; RECODER, Maria-José; ABADAL, Ernest. **Informação eletrônica e novas tecnologias**. Tradução de Dinah Aguiar Población. São Paulo: Summus Editorial, 1995. (Coleção Novas Buscas em Comunicação).
- SMITH, Anthony. **Goodbye Gutenberg: The newspaper revolution of the 1980s**. New York: Oxford University Press, 1980.
- SOL, Selena. **Introduction to databases for the web: Pt.1**. 1998. In: [www.databasejournal.com/sqletc/article.php/1428721](http://www.databasejournal.com/sqletc/article.php/1428721). Acesso em: 19/09/2006.
- SOUSA, Artur Afonso. **Bases de dados Web e XML**. Lisboa: FCA Editora de Informática, 2002.

- SPIVACK, Nova. **Towards a World Wide Database (WWDB)**. 2005. Disponível em: [http://novaspivack.typepad.com/nova\\_spivacks\\_weblog/2005/10/towards\\_a\\_world.html](http://novaspivack.typepad.com/nova_spivacks_weblog/2005/10/towards_a_world.html). Acesso em: 02/12/2006.
- SQUIRRA, Sebastião. **Jorn@lismo Online**. São Paulo: Arte & Ciência, 1998.
- STOCKINGER, Gottfried. **A sociedade da comunicação**. O contributo de Niklas Luhman. Rio de Janeiro: Papel & Virtual, 2003.
- SULEMAN, Din. **Google wants to be part of journalism's future**. 2002. Disponível em: <http://www.rediff.com/us/2002/oct/18us.htm>. Acesso em: 05/12/2006.
- SUMMIT, Roger. **Reflections on the Beginnings of Dialog**. Birth of Online Information Access. 2002. Disponível em: <http://support.dialog.com/publications/chronolog/200206/1020628.shtml>. Acesso em: 27/04/2005.
- THE LEXISNEXIS TIMELINE. CELEBRATING INNOVATION . . . AND 30 YEARS OF ONLINE LEGAL RESEARCH. 1973-2003**. Disponível em: [www.lexisnexis.com](http://www.lexisnexis.com). Acesso em: 28/05/2005.
- THOMAS, Gérard. **La presse régionale parie sur le web de proximité**. In: *Libération*. Disponível em: <http://www.liberation.fr/actualite/medias/194398.FR.php>. Acesso em: 23/07/2006.
- THOMAS, Lou. **'Website must come second' says Independent's Kelner**. In: *Online press gazette*. Julho/2006. Disponível em: [http://www.pressgazette.co.uk/article/200706/independent\\_website\\_second](http://www.pressgazette.co.uk/article/200706/independent_website_second). Acesso em 23/07/2006.
- TRAMULLAS y KRONOS, Jesus. **Las bases de datos multimédia**. In: *Introducción a la documática*. (1997-2000). Disponível em: <http://tramullas.com/documatica/6.html>. Acesso em: 29/08/2006
- TRAQUINA, Nelson. **O que é jornalismo**. Lisboa: Quimera, 2002.
- TRIVIÑOS, Augusto N.S. **Introdução à pesquisa em ciências sociais**. A pesquisa qualitativa em educação. São Paulo: Atlas, 1987.
- VARELA, Juan. **Periodismo participativo: El periodismo 3.0**. In: ORDUÑA, Octavio I. R. et al. **Blogs - La conversación en Internet que está revolucionando medios, empresas y a ciudadanos**. Madrid: ESIC, 2005. pp: 77-163.
- \_\_\_\_\_. **Política 3.0**. In: *Periodistas 21*. 24 de março, 2004. Disponível em: <http://periodistas21.blogspot.com/2004/03/politica-3.html>. Acesso em: 29/08/2005.

VELLOSO, Fernando de Castro. **Informática** – Conceitos básicos. Rio de Janeiro: Campus, 1999.

VESNA, Victoria. **Database aesthetics: of containers, chronofiles, time capsules, xanadu, Alexandria and world brain.** 1999. Disponível em: [http://vv.arts.ucla.edu/publications/publications/98-99/ai\\_society/vesna\\_essay.html](http://vv.arts.ucla.edu/publications/publications/98-99/ai_society/vesna_essay.html). Acesso em: 27/07/2006.

\_\_\_\_\_. **Database are us.** In: *AI Soc.* 14, 2 (May. 2000a), 157-175. Disponível em: <http://dx.doi.org/10.1007/BF01205446>. Também disponível em: <http://www.springerlink.com/media/49te8hwxqndjv6tvxceg/contributions/g/1/8/7/g187821393135271.pdf>. Acesso em 27/07/2006.

\_\_\_\_\_. **Database aesthetics.** In: *AI Soc.* 14, 2 (May. 2000b), 155-156. Disponível em: <http://dx.doi.org/10.1007/BF01205445>. Acesso em: 27/07/2006.

VESNA, Victoria; Gil, K.; SMITH, D.; VAUX, J. (editors). **Database aesthetics: Issues of organization and category in online art.** In: *AI Society: The Journal of Human-Centred and Machine Intelligence.* Springer. Fab/March/1999. Disponível em: [http://vv.arts.ucla.edu/publications/publications/98-99/ai\\_society/index.html](http://vv.arts.ucla.edu/publications/publications/98-99/ai_society/index.html). Acesso em: 27/07/2006.

**VII SEMANA DO AUDIOVISUAL E MULTIMÉDIA.** Universidade Lusófona de Humanidades e Tecnologia (ULHT). Lisboa, Maio, 2005

VIANNA, Ruth Penha Alves. **Informatização da imprensa brasileira.** São Paulo: Loyola, 1992.

VIERA, Lise. **L'édition électronique. De l'imprimé au numérique: Évolutions et Stratégies.** (L@byrinthes). Pessac: Presses Universitaires de Boudeaux, 2004.

VILLEGAS, P. y GARCÍA, D.: **Sistemas de descripción de contenidos multimedia.** In: *Comunicaciones de Telefónica I+D*, núm. 24, Enero, Telefónica Investigación y Desarrollo, Madrid, 2002, pág. 133-144. Disponível em: <http://www.tid.es/presencia/publicaciones/comsid/esp/24/art7.pdf>, e também em: <http://www.telefonica.es/convergenciademedios/documentosdeinteres/pdf/art71.pdf>. Acesso em 29/08/2006.

VIVAR, Jesús Flores; ARRUTI, Alberto Miguel. **Gestión del conocimiento en los médios de comunicación.** Madrid: Fraga Editorial, 2004.

WARD, Mike. **Journalism Online.** London: Focal Press, 2002.

WARNER, William B. **Computable Culture and the Closure of the Media Paradigm: Lev Manovich's The Language of New Media.** Disponível em: [http://www.manovich.net/LNM/review\\_Warner.doc](http://www.manovich.net/LNM/review_Warner.doc). Acesso em: 10/09/2006.

- WATTERS, C.R.; et al. **An Evaluation of Two Metaphors for Electronic News Presentation**. 1997, 2000. In: 8th International Conference on Digital Documents and Electronic Publishing, DDEP 2000; 5th International Workshop on the Principles of Digital Document Processing, PODDP. Munique/Alemanha, 2000. Também disponível em: <http://users.cs.dal.ca/~watters/watters.doc>. Acesso em: 14/12/2006.
- WEBSTER, Frank. **The information society revisited**. In: LIEVROUW, Leah A.; LIVINGSTONE, Sonia. **Handbook of new media**. Social Shaping and Consequences of ICTs. London: Sage, 2002.
- WILLIAMS, Raymond. **The television as technological form**. London: Routledge, 1990.
- \_\_\_\_\_. **From medium to social practice**. In: **Marxism and literature**. London: Oxford University Press, 1977.
- WOLF, Mauro. **Teorias da comunicação**. Lisboa: Presença, 1992.
- WOLTON, Dominique. **E depois da internet?** Algés – Portugal, Difel, 2000. Título original: Internet et après?
- WRIGHT, Peggy. **Knowledge Discovery In Databases: Tools and Techniques**. (1998). Disponível em: <http://www.acm.org/crossroads/xrds5-2/kdd.html>. Acesso em: 28/11/2005.
- YIN, Robert K. **Estudo de caso**. Planejamento e métodos. Tradução Daniel Grassi. (3ª edição). Porto Alegre: Bookman, 2005.
- \_\_\_\_\_. **Case Study Research: design and methods**. EUA: Sage Publications, second edition, 1994.
- ZAMORA, Lizy Navarro. **Los periódicos on line**. San Luis de Potosí: Editorial Universitaria Potosina, 2002.
- ZEIGLER, Todd. **9 Ways for Newspapers to Improve Their Websites**. 2006. Disponível em: <http://www.bivingsreport.com/2006/9-ways-for-newspapers-to-improve-their-websites>. Acesso em: 04/10/2006.

#### **Outras fontes de pesquisa :**

- American Journalism Review: <http://www.ajr.org/index.asp>.
- Associação Brasileira de Jornalismo Investigativo: <http://www.abraji.org.br/>.
- Blackwell Synergy: <http://www.blackwell-synergy.com/>.
- Biblioteca do Conhecimento Online: <http://www.b-on.pt/>.

- Biblioteca On-line de Ciências da Comunicação (*Bocc*): <http://www.bocc.ubi.pt/>.
- Cambridge Journals Online: <http://journals.cambridge.org/>.
- Compós: <http://www.compos.org.br/>.
- Encyclopedia Britannica Online: <http://www.britannica.com/>.
- Emerald: <http://hermia.emeraldinsight.com/Insight/menuNavigation.do?hdAction=InsightHome>.
- FOLDOC (Free On-line Dictionary of Computing): <http://foldoc.org/>.
- Grupo de Pesquisa em Jornalismo On-line (GJOL): <http://www.facom.ufba.br/jol>.
- HW Wilson: <http://vnweb.hwwilsonweb.com/>.
- IBICT - Instituto Brasileiro de Informação em Ciência e Tecnologia:  
<http://www.ibict.br/>.
- Intercom: <http://www.intercom.org.br/index.shtml>.
- Investigative Reporters and Editor's (IRE): <http://www.ire.org/>.
- MMI - Media Monitor Internacional: <http://www.mmi-e.com/>.
- Online Journalism Review: <http://www.ojr.org/>.
- Oxford Journals: <http://www.oxfordjournals.org/>.
- Portal Periódicos/CAPES (O Portal Brasileiro da Informação Científica):  
<http://www.periodicos.capes.gov.br/portugues/index.jsp>.
- SAGE Publications: <http://online.sagepub.com/>.
- SBPJor: <http://www.sbpjor.ufsc.br/>.
- Scielo Brasil (*Scientific Electronic Library Online*):  
[http://www.scielo.br/scielo.php/script\\_sci\\_home/Ing\\_pt/nrm\\_iso](http://www.scielo.br/scielo.php/script_sci_home/Ing_pt/nrm_iso).
- Sociedade Brasileira de Computação: <http://www.sbc.org.br/>.
- The ACM Portal: <http://www.acm.org>.
- The Institute for Interactive Journalism: <http://www.j-lab.org/>.
- Telecommunications: Glossary of Telecommunication Terms:  
<http://ntia.its.bldrdoc.gov/fs-1037/fs-1037c.htm>.
- The Online News Association: <http://www.journalist.org/>.
- The Third Culture - Contents: <http://www.edge.org/documents/ThirdCulture/d-Contents.html>.
- WhatIs?com - *IT encyclopedia and learning center*: <http://whatis.techtarget.com/>.
- Wikipedia- *The free encyclopedia*: [http://en.wikipedia.org/wiki/Main\\_Page](http://en.wikipedia.org/wiki/Main_Page).

## GLOSSÁRIO

**Agregadores de notícia** - *Sites* que funcionam, em geral, de modo inteiramente automatizado, com algoritmos e tecnologias específicas, técnicas de inteligência artificial, entre outras, associadas às bases de dados. Produzem sínteses, resumos, a partir da indexação de informações produzidas por outros *sites* jornalísticos, também remetendo às fontes originais para a leitura dos textos na íntegra.

**Ajax** - *Asynchronous JavaScript and XML*. Termo que se refere ao uso de um grupo de tecnologias para criar aplicações *web* interativas. Ela proporciona às páginas *web* trocarem pequenas quantidades de dados com o servidor e, assim, essas páginas não têm que ser recarregadas cada vez que o usuário fizer alguma mudança.

**Alert news** – Serviço oferecido por diversos *sites*, especialmente os jornalísticos, para notificar aos usuários cadastrados em suas bases de dados sobre novas informações previamente solicitadas. No ciberespaço, segundo Machado (2006), constitui numa importante ferramenta para a apuração das informações.

**ASP** – *Active Server Page*. Ambiente de programação desenvolvido pela *Microsoft* para elaboração de páginas *web* dinâmicas. Ela mescla *HTML* com outras linguagens de programação. Seu principal concorrente é a *PHP*.

**Algoritmo** – Na Matemática e nas Ciências da Computação, significa um pequeno procedimento para resolver um problema recorrente. Ele representa os passos necessários para se realizar uma tarefa e a sua implementação pode ser feita por um computador ou por um ser humano. Diferentes algoritmos podem realizar a mesma tarefa usando um conjunto diferenciado de instruções em mais ou menos tempo, espaço ou esforço do que outros. Um programa de computador pode ser considerado como um algoritmo explícito.

**Arquitetura da informação** – Esquema abstrato dos conteúdos de um produto jornalístico digital plasmado em uma estrutura de base de dados, promovendo a simbiose entre o servidor, o usuário e o sistema de gestão de conteúdos. Permite partir de uma base previamente programada para gerar um *site* dinâmico. Cumpre as funções de: mapa, e orientação para busca e recuperação (LÓPEZ; GAGO; PEREIRA, 2003). É concebida como um roteiro audiovisual que baliza a produção. Consiste na explicação tela a tela dos possíveis desdobramentos da matéria, tendo o cuidado com a linha narrativa (SCHWINGEL, 2004, 2005).

**Back end** – Se refere à estrutura em si de uma base de dados. É formado pelos algoritmos e conjuntos de dados que ficam escondidos, isto é, não são visíveis aos usuários finais do *site*. São acessíveis aos administradores de um sistema de base de dados.

**Banda larga** – Redes com capacidade para transmitir grandes volumes de informações em áudio e em vídeo em alta velocidade.

**Base de dados (BD)** – Do inglês *database*. O termo tem sido usado desde a metade dos anos 60. Denomina qualquer grande coleção integrada de dados (texto, números, imagens fixas, e objetos multimídia, como vídeo, infografias interativas, *slideshows*, etc.), arquivados de forma estruturada. É um sistema de armazenamento de dados baseado em computador, cujo objetivo global é registrar e manter informação. Esta informação pode ser qualquer uma considerada significativa à organização servida pelo sistema. Qualquer coleção de dados ou informações que sejam organizadas para rápida pesquisa e recuperação por um computador é uma base de dados. As BDs são estruturadas para facilitar o armazenamento, a recuperação, modificação e manutenção dos dados, conforme várias operações de processamento de dados. De acordo com Paul (2004), um dos traços distintivos entre as bases de dados digitais e os seus predecessores analógicos é a sua possibilidade inerente para a recuperação e filtragem dos dados (ou dos registros nelas inseridos) em múltiplas maneiras. Uma BD constitui o hardware que armazena o dado; o software que permite alojar o dado em seu *container* respectivo e recuperá-lo, filtrá-lo e mudá-lo, tanto como os usuários que acrescentam um nível melhor no entendimento do dado como informação (PAUL, 2004). Há alguns autores, como Colle (2002), que diferenciam “base”, como sendo a estrutura lógico-matemática, de “banco”, relacionado ao conteúdo, à informação. Em geral, nas Ciências da Computação, o termo base de dados é empregado para denominar ambas as esferas. Coube a Manovich (2001) ampliar o significado das BDs na contemporaneidade, reconhecendo-as como uma forma cultural simbólica e percebendo-as como lógica e estrutura para forjar boa parte dos objetos da nova mídia.

**Blogsfera** – É o universo e a cultura dos *weblogs*. Trata-se de um sistema complexo, auto-regulado, dinâmico e também sensível à informação produzida pelos meios de informação tradicionais. Por meio dos *links* remetendo para *sites* externos e para *blogs* diversos, cada “blogueiro” se insere dentro de uma comunidade conformada por vínculos comuns e pela familiaridade que adquire com seu grupo de referência (ORIHUELA, 2005).

**Broadsheet metaphor** – Metáfora do impresso. Parâmetro utilizado para a organização, a hierarquização e a apresentação das informações em *web sites* jornalísticos de modo semelhante ao jornal impresso. Por exemplo: denominação das seções, múltiplas colunas, múltiplas histórias com variados tamanhos para os destaques e fotografias e vídeos integrados.

**Computer Assisted Reporting** – De acordo com Garrison (1998), a *CAR* ou Reportagem Assistida por Computador, é a aplicação de computadores para obtenção e análise de informações para a elaboração de matérias jornalísticas. Envolve o uso de computadores de todos os tamanhos, de *mainframes* a PCs até os computadores de mão. Também chamada *Computer Assisted Journalism* ou Jornalismo Assistido por Computador. Para Paul (1999) foram as bases de dados que fundaram a Reportagem Assistida por Computador.

**Content Management Systems** – Sistemas de Gerenciamento de Conteúdo. São ferramentas para publicação de conteúdos de qualquer natureza na WWW. Os

*blogs*, os conhecidos diários da rede, são um exemplo de ferramenta de publicação.

**Cascading Style Sheets** – Folhas de Estilo em Cascata. Permitem que seja retirada da marcação *HTML* toda a formatação (apresentação) do documento *web*. Determina cores, formas, tipos e tamanhos, posicionamentos em todo o visual da página.

**Ciberespaço** – Espaço conceitual em que os usuários interagem através de tecnologias de comunicação mediadas pelo computador. Termo cunhado pelo escritor canadense William Gibson no livro *Neuromancer* (1984), considerado o marco inaugural do gênero cyberpunk, o primeiro movimento cultural a refletir criticamente sobre as implicações trazidas pelas novas tecnologias da informática.

**Cluster** – Também chamado *clustering* (quando visto em um contexto mais abrangente). Cluster pode ser definido como um sistema onde dois ou mais computadores trabalham de maneira conjunta para realizar processamento pesado. De acordo com Colle (2002), indica, em estatísticas, uma ação de reagrupamento de dados em função de um ou vários critérios de classificação ou seleção. Em informática, implica que duas ou mais tabelas de uma base de dados fique guardada fisicamente de forma adjacente para obter um melhor rendimento quando se precisa consultá-las simultaneamente.

**Clustering algorithm** – É um procedimento matemático que encontra similaridades entre os elementos, e agrupa-os, conforme especificação prévia, por categorias, seções, etc. *Clustering* quer dizer segmentação e consiste em identificar grupos de registros que são similares entre si, mas diferentes do resto dos dados e, especialmente, as variáveis que provêm o melhor agrupamento possível. Alguns *sites* agregadores de notícia que funcionam de forma totalmente automatizada trabalham com este procedimento. Ele permite examinar artigos de diferentes fontes, analisando fatores como: o assunto tratado no artigo, as pessoas e as organizações cujos nomes são mencionados juntamente com o da primeira pessoa, como um mesmo assunto foi tratado em outros idiomas, etc. Funciona como uma biblioteca ou serviço de *clipping*, descobrindo, combinando e colecionando artigos.

**Crawlers** - Robôs que buscam informações em *sites* noticiosos. É uma metodologia empregada pelo *Google News*, pelo *Topix.net*, entre outros agregadores.

**Database aesthetics** – “Estética base de dados”. Termo usado para descrever os princípios estéticos aplicados na imponente lógica das bases de dados para qualquer tipo de informação, filtrando coleções de dados, e visualizando os dados.

**Data mining** – Mineração de dados. Exploração de dados. Conjunto de operações de diversos tipos que se pode realizar para explorar relações, tendências e projeções em conjuntos de dados de grandes dimensões (COLLE, 2002). As ferramentas de *data mining* se baseiam no conhecimento acumulado que está invisível nas BDs. Forma automática de aquisição de conhecimento. Uma técnica que torna um sistema inteligente à medida que auxilia no processo de decisão sem a interferência humana. Trata-se de um processo não-trivial de identificar, em

dados, padrões válidos, potencialmente úteis. Navega (2002) ressalta que o *data mining* ainda requer uma interação muito forte com os analistas humanos, pois, em última instância, são eles os principais responsáveis pela determinação do valor dos padrões encontrados.

**Fractal** – Termo cunhado em 1975 por Benoit Mandelbrot para descrever um objeto geométrico que nunca perde a sua estrutura qualquer que seja a distância de visão. É uma forma geométrica, de aspecto irregular ou fragmentado, que pode ser subdividida indefinidamente em partes, as quais, de certa maneira, são cópias reduzidas do todo. A palavra fractal significa auto-semelhante. Auto-semelhança é a simetria através das escalas. Em 1980, Mandelbrot descobriu o primeiro fractal gerado por computador.

**Front end** – Presença gráfica do sistema. Parte visível experienciada pelos usuários, por meio das interfaces. A estética base de dados está relacionada com o *front end*, e com as suas manifestações visuais e implicações culturais.

**Hipertexto** – Conforme definido por Theodor Nelson, em 1974, é a escrita não-sequencial, num texto que se bifurca, que permite que o leitor escolha e que se leia melhor numa tela interativa. De acordo com a noção popular, trata-se de uma série de blocos de textos conectados entre si por nexos (*links*), que formam diferentes itinerários para o usuário.

**Hipermídia** – Definida por Leão (1999) como uma tecnologia que engloba recursos do hipertexto e multimídia, permitindo ao usuário a navegação por diversas partes de um aplicativo, na ordem que desejar. Gosciola (2003) a define como um conjunto de meios que permite acesso simultâneo a textos, imagens e sons de modo interativo e não-linear, possibilitando fazer *links* entre elementos de mídia, controlar a própria navegação e, até, extrair textos, imagens e sons cuja seqüência constituirá uma versão pessoal desenvolvida pelo usuário.

**Hipernarrativa** – É definida por Lev Manovich (2001) como aquela que resulta da soma de múltiplas trajetórias efetuadas através das bases de dados. O usuário da hipernarrativa atravessa as BDs seguindo os *links* entre os registros, conforme estabelecido pelo criador da base de dados. Assim, a narrativa é construída pela linkagem entre os elementos de uma base de dados em uma ordem particular, que define uma trajetória conduzindo elementos de um ponto para outro.

**HTML** - *Hypertext Markup Language* ou Linguagem de Marcação em Hipertexto. Linguagem que define a forma de exibição dos conteúdos na página *web*.

**Information Visualization** - Visualização da Informação. Representa uma tendência estabelecida no moderno *design* de interface para o usuário (SHNEIDERMAN, 2004). Trata-se de um campo de pesquisa distintivo, que tem se expandido rapidamente e se tornado um campo interdisciplinar de pesquisa.

**Interface** – Numa definição simples, interface é o todo da tela que se apresenta diante do usuário e que é denominada habitualmente como *GUI (Graphical User Interface)*. É através dela que se interage com o conteúdo (textos, fotografias,

vídeos, áudios, filmes, músicas, infografias, etc.) dos *sites* e dos ambientes virtuais.

**Internet** – Conjunto das redes de computadores interligados que funcionam segundo o protocolo TCP/IP (*Transmission Control Protocol/Internet Protocol*) e que permite o acesso a informações e todo o tipo de transferência de dados.

**Intranet** – Sistema interconectado das redes locais de computadores das empresas.

**Java** – Linguagem de programação desenvolvida pela *SUN Microsystems* com a propriedade de poder ser utilizada de forma independente do sistema operacional (*UNIX, Mas OS, Windows, Linux*).

**Jezebel's Tomb** - Novela baseada em livro homônimo escrito pelo repórter David Hilzenrath da seção “Business” do jornal *The Washington Post*. É a primeira vez que o *site Washingtonpost.com* publica conteúdo ficcional. A novela estreou no dia 17 de janeiro de 2007.

**Jornalismo digital** – Modalidade jornalística em desenvolvimento há mais de uma década. Segundo Machado (2006), é o jornalismo em que todas as etapas do processo de produção e circulação de conteúdos - apuração, composição, circulação e modelo de negócios - são desenvolvidas no ciberespaço. Engloba os produtos jornalísticos na web, bem como os recursos e tecnologias disponíveis utilizados para a disseminação dessa produção também para dispositivos móveis, como celulares, *iPods, MP3 players*, computadores de mão, *smartphones*, etc.

**Jornalismo Digital em Base de Dados (JDBD)** – É o modelo que tem as bases de dados como definidoras da estrutura e da organização, bem como da apresentação dos conteúdos de natureza jornalística, de acordo com funcionalidades e categorias específicas, que vão permitir a criação, a manutenção, a atualização, a disponibilização e a circulação de produtos jornalísticos digitais dinâmicos. Trata-se de um paradigma na transição entre a terceira e uma quarta geração do jornalismo digital, com potencial para fazer emergir uma nova metáfora para a apresentação dos conteúdos jornalísticos.

**Jornalismo participativo/Jornalismo 3.0** – Segundo Varela (2005), é a terceira versão do jornalismo digital. A 1.0 é o jornalismo que transpõe o conteúdo tradicional de meios analógicos para o ciberespaço; a 2.0 é a criação de conteúdo de e para a rede e a 3.0 é aquela que socializa este conteúdo e os próprios meios. Os dois lemas deste jornalismo participativo são: o público sabe mais das notícias e das informações que os próprios jornalistas (ponto defendido por Gilmor, 2004, 2005) e que a informação deve ser uma conversação de muitos para muitos. Assim, o jornalismo participativo constrói a informação a partir da conversação, onde a participação dos usuários é fundamental para concluir o discurso e a informação.

**KDD** – *Knowledge Discovery in Database*. Descoberta de Conhecimento em Bases de Dados. Instrumento para a prospecção de conhecimento embutido nos dados. Ramo da Computação surgido no final dos anos 80 com o objetivo de encontrar uma maneira automatizada de explorar as bases de dados e reconhecer os padrões

existentes através da modelagem de fenômenos do mundo real. O KDD refere-se às etapas que produzem conhecimentos a partir de dados relacionados e sua principal característica é a extração não-trivial de informações implicitamente contidas em uma base de dados. Essas informações são de difícil detecção por métodos tradicionais de análise e devem ser potencialmente úteis para tomada de decisão. Enquanto os métodos tradicionais são capazes de tratar apenas as informações explícitas, o KDD é capaz de detectar informações implícitas armazenadas nas BDs.

**Lexia** – Blocos de informações.

**Link** – De *hiperlink*. É o vínculo eletrônico que conecta os documentos, textos, matérias, artigos, objetos multimídia. É uma referência ou elemento de navegação que traz automaticamente a informação respectiva para o usuário quando um elemento é selecionado por ele enquanto navega em um *site*.

**Meios sociais** – Se definem pela convergência de indivíduos em redes sociais, o uso de novos meios e a sindicância ou vinculação de idéias, escritos e outros conteúdos informativos e de opinião (VARELA, 2005).

**Metadados** – Matéria-prima das bases de dados. São dados acerca dos dados ou informação acerca da informação. Contêm informação semântica sobre os dados. São ferramentas que auxiliam tanto na identificação de informação pontual, como para extrair informação sobre o conjunto e que provêem um contexto que pode ser de grande importância para uma melhor interpretação e análise das informações. O RSS ou sindicância, recurso de descrição de conteúdos, é uma das conhecidas aplicações que utilizam os metadados.

**Metainformação** – Informação acerca de um conjunto de informações. Segundo Colle (2002), este conceito se utiliza essencialmente para designar conhecimentos associados a conjuntos de dados e, no campo da Informática, remetem em particular ao conhecimento relacionado com a estrutura e os conteúdos das bases de dados.

**MySQL** – Software livre, cujo desenvolvimento foi baseado na linguagem *SQL*. É uma das bases de dados mais populares atualmente.

**Ontologias** - Conhecimento sobre relações entre palavras, dados, objetos. Uma ontologia é um vocabulário controlado que descreve objetos e as relações entre eles em um modo formal. São especificações de conceitualizações utilizadas para o compartilhamento do conhecimento.

**PEAR** - *PHP Extension and Application Repository*. É um *framework* e sistema de distribuição para reuso de componentes em PHP. A idéia é reutilizar o que já está pronto. Os códigos são chamados de pacotes

**PHP** - *PHP: Hypertext Preprocessor*. Ambiente de programação desenvolvido pela comunidade do software livre. Trata-se de uma linguagem de programação de computadores interpretada, livre e muito utilizada para gerar conteúdo dinâmico na WWW. Apesar de ser uma linguagem de fácil aprendizado e de uso para

pequenos *scripts* dinâmicos simples, o PHP é uma linguagem poderosa orientada a objetos.

**Podcasting** – Técnica digital que permite importar da internet um conteúdo de áudio e/ou de vídeo para ouvir/vê-los em *iPods*, por exemplo. Cada usuário pode importar para o seu computador os conteúdos disponíveis, seja uma entrevista, um vídeo de um acontecimento ou outro tipo de conteúdo disponível. Para isso, é necessário instalar um programa como o *iTunes* (gratuito), com versões para *PC* e *Mac*, que funcionará como receptor de *podcasts*.

**Remediation** - É a representação de um meio em outro, implicando no reconhecimento do meio anterior, da sua linguagem e da sua representação social para se estabelecer em um novo suporte. Com este conceito, Bolter & Grusin (2000) querem significar que todos os meios passam por processos de remediação, e melhorias, uma vez que têm o seu sistema de produção afetado pela chamada nova mídia.

**Resolução semântica** – Conceito introduzido por António Fidalgo (2003, 2004). Conforme ele explica, como uma imagem digital aumenta a sua qualidade com o aumento da resolução gráfica (o número de *pixels* por centímetro quadrado), também a pluralidade e a diversidade das notícias *on-line* sobre um evento aumentam a informação sobre o mesmo, aumentando, assim, a resolução semântica. O autor afirma que o conceito é adequado para explicar a sucessão de notícias na informação *on-line* dada em contínuo e só faz sentido se os elementos informativos sucessivos e progressivos seguirem uma ordenação. Inclui as noções de baixa e alta resolução.

**RSS** - Abreviação de *Really Simple Syndication* (Distribuição Realmente Simples). É uma maneira de distribuir informação por meio da internet que se tornou uma poderosa combinação de tecnologias "pull" - com as quais o usuário da web solicita as informações que deseja - e tecnologias "push" - com as quais informações são enviadas a um usuário automaticamente. O visitante de um *site* que funcione com *RSS* pode solicitar que as atualizações lhe sejam enviadas (processo conhecido como "assinando um feed"). Através da tecnologia *RSS*, os *sites* jornalísticos geram um fluxo automatizado de notícias escrito na linguagem *XML* que podem ser incorporadas a qualquer produto *web* preparado para tal, em um processo que se chama *sindicação*. As notícias também podem ser distribuídas para dispositivos móveis, como os celulares.

**SGBD** - Sistema de Gerenciamento de Base de Dados. Coleção de programas que permite a criação e a manutenção de BDs.

**Sistemas de gestão de conteúdo** – Também chamados sistemas publicadores. São ambientes tecnológicos que possuem as potencialidades dos sistemas de administração de conteúdos e de gerenciamento de portais, adaptados às rotinas produtivas do jornalismo digital (SCHWINGEL, 2004; 2005). As bases de dados são as balizadoras de tais sistemas.

**Sistemas inteligentes** – Um tipo de sistema que aprende a responder de forma automática às demandas do seu entorno, possibilitando alcançar os objetivos previamente estipulados pelo programador. O jornalismo digital, diz Machado (2006), depende da criação de sistemas inteligentes ao longo de todo o processo de produção.

**Site estático** – Produto único, ainda que recorrendo a *templates* (modelos). É constituído por páginas estáticas, sempre com o mesmo conteúdo. Neste tipo de *site* construído apenas em *HTML*, para representar visualmente 5.000 notícias, são necessárias exatamente 5.000 páginas diferentes.

**Site dinâmico** – Produto estruturado em base de dados. O conteúdo é dinâmico e será sempre o resultado de uma pesquisa (*query*) nas notícias inseridas e da estrutura da BD. Os *web sites* dinâmicos contam com uma, duas, ou três páginas físicas para representar visualmente, por exemplo, 10 ou 15 mil notícias, pois são flexíveis e a informação pode ser representada a partir de um sistema que negocia com o servidor e o usuário (LÓPEZ, GAGO, PEREIRA, 2003). Outra diferença é que este tipo de *site* utiliza linguagem de programação, o que permite a modificação automática, sem a intervenção manual de um programador (MIELNICZUK; MARQUES, 2006).

**Sites sociais** – Conjunto formado por *sites* de conteúdos colaborativo (como *Digg*), *weblogs*, as *wikis*, redes sociais, fóruns, programas de mensagem instantânea, grupos de notícia, *sites* de compartilhamento de fotos e de vídeos, listas de distribuição de e-mail, entre outros. Chamados meios sociais por empregarem ferramentas (programas e redes) que permitem ou facilitam a interação de grupos geograficamente dispersos (ORDUÑA et al., 2005).

**Slideshows** – Formato específico para apresentar conteúdos digitais, combinando texto, imagens e sons. Podem ser usados para apresentar desde um simples ensaio fotográfico sobre assunto específico, para conformar uma sequência de imagens com narração até para empacotar reportagens multimídias.

**SMS** – *Short Message System*. Sistema de comunicação com base num texto, amplamente utilizado em celulares. Usando o *SMS*, os usuários podem receber notícias, entre outras mensagens enviadas por bancos, e trocar entre si curtas mensagens escritas.

**Software livre ou de fonte aberta** – Software que é desenvolvido, mantido e distribuído coletivamente a muitas pessoas de todo o mundo. O software de fonte aberta está à disposição de todo e é “aberto”, permitindo a alteração do código fonte. Exemplos são o sistema operacional *Linux*, o software do servidor *Apache* e o software de base de dados *MySQL*.

**SQL** – *Structured Query Language* ou Linguagem de Consulta Estruturada, da IBM, que se tornou o padrão dos SGBDs relacionais. Utilizada para comunicar com sistemas de bases de dados.

**Tags** – Estruturas de linguagem de marcação que consistem em breves instruções, tendo uma marca de início e outra de fim.

**Templates** – Modelos prontos.

**TrackBack** - Um tipo de *LinkBack*, que permite que se receba notificação quando alguém coloca um *link* para determinado artigo ou outro material contido no *site*. Também assegura à empresa proprietária de um *site* ou autor de um *blog* saber quem está linkando ou remetendo para seu conteúdo.

**Tree map** – Técnica criada por Ben Shneiderman para a visualização de informação hierarquizada, que gera o padrão de visualização conhecido como *Squarified* (retângulos alongados). Vem sendo empregada por *sites* jornalísticos como o *Newsmap*.

**Videotexto** - Sistema de transmissão de informações por cabo ou sem fio que funciona por meio de um computador pessoal ou terminal especial, ou de aparelho de televisão conectado à rede telefônica por um *modem* ou aparelho decodificador de sinais analógicos em digitais e vice-versa.

**WAP** – *Wireless Application Protocol* ou *Protocolo de Aplicações Sem Fio*. É um padrão internacional para aplicações que utilizam comunicações sem fio (internet móvel), como por exemplo acesso à web ou ao correio eletrônico a partir de celulares, entre outros dispositivos móveis.

**Web 2.0** - Segundo Tim O’Reilly (2005), que cunhou o termo, é um conjunto de princípios e práticas que formam uma espécie de núcleo gravitacional em torno da ideia de Web 2.0, proposta como a segunda geração da internet, também chamada “Web social”. A disponibilização de aplicações e serviços para compartilhamento do conhecimento e da produção social e livre de conteúdos é um dos eixos da Web 2.0. Para o jornalismo, a Web 2.0 configura o que alguns autores denominam como Jornalismo 3.0, ou jornalismo participativo (VARELA, 2005).

**Web 3.0** – Chamada *Semantic Web* ou *Web Semântica*. É uma evolução da World Wide Web, na qual a informação é processável na máquina, melhor do que sendo apenas orientada aos humanos. Permite, deste modo, usar agentes de software para encontrar, compartilhar e combinar informações visando uma utilização mais fácil por parte dos usuários.

**Weblog** – Ou *blog*. É uma página da Web cujas atualizações (chamadas *posts*) são organizadas cronologicamente (como um histórico ou diário). No jornalismo, são usados como um formato tanto por jornalistas de modo independente, como por *sites* das empresas do *mainstream*, que a cada dia ampliam o espaço para *blogs* de seus jornalistas, além de agregar os de colaboradores e de usuários. É um dos aspectos principais do jornalismo cidadão ou jornalismo participativo.

**Wi-fi** - Abreviatura para *wireless fidelity* (marca registrada pertencente à *Wireless Ethernet Compatibility Alliance -WECA*). É uma tecnologia de interconexão entre dispositivos sem fio, usando o protocolo IEEE 802.11b. O padrão *Wi-Fi* opera em faixas de frequência que não necessitam de licença para instalação e/ou operação. No entanto, para uso comercial no Brasil é necessária licença da Agência Nacional de Telecomunicações (Anatel).

**Wiki** – Software que permite a qualquer pessoa, utilizando um motor de busca, criar e editar páginas na Web. A *Wikipedia* ([www.wikipedia.org](http://www.wikipedia.org)), enciclopédia *on-line*, é um exemplo de um wiki, assim como o *Wikinews* ([http://en.wikinews.org/wiki/Main\\_Page](http://en.wikinews.org/wiki/Main_Page)), *site* de notícias com versões em mais de 200 idiomas, incluindo o português ([http://pt.wikinews.org/wiki/P%C3%A1gina\\_principal](http://pt.wikinews.org/wiki/P%C3%A1gina_principal)) no qual qualquer pessoa pode publicar conteúdos (texto, áudio, vídeo, fotografias).

**World Wide Database (WWDB)** – Denominação proposta por Nova Spivak (2005). É uma rede globalmente distribuída de registros de dados que residem em milhões de nós ao redor da rede e que se comporta coletivamente como um gigante virtual, um sistema descentralizado de base de dados. Seria o nível 4 de evolução para a WWW, e combinaria o nível 3 (web semântica), com o nível 2 (web de dados) e o nível 1 (web de documentos). Dentro desta rede, serviços agregadores, *remix*, e subconjuntos de dados dentro de BDs virtuais sobre vários assuntos podem ser referenciados em múltiplas e distintas aplicações de contextos.

**World Wide Web (WWW)** – Face gráfica e hipertextual da internet, dedicada à transmissão de informação multimídia.

**XML** – *eXtensible Markup Language*. É uma (meta) linguagem de marcação de documentos completamente independente das plataformas de hardware e software que utilizam, sendo um padrão aberto. É uma linguagem de estruturação de dados que habilita a produção de documentos no qual o dado é separado da forma ou da apresentação dele. Padrão recomendado pelo *W3C – World Wide Web Consortium*, que é liderado por Tim Berners-Lee.

## APÊNDICES

## APÊNDICE I

### PRODUTOS QUE COMPÕEM O *CORPUS* EMPÍRICO

- 1- *Akademia*: [www.labcom.ubi.pt](http://www.labcom.ubi.pt)
- 2- *BBC News*: <http://news.bbc.co.uk>
- 3- *Blog Deu no jornal*: <http://deunojornal.zip.net/>
- 4- *El Mundo*: [www.elmundo.es](http://www.elmundo.es)
- 5- *El País*: [www.elpais.com](http://www.elpais.com)
- 6- *Folha Online* : <http://www.folha.uol.com.br/>
- 7- *Google News*: <http://news.google.com/?hl=en>
- 8- *International Herald Tribune*: [www.iht.com](http://www.iht.com)
- 9- *Libération*: <http://www.liberation.fr/>
- 10- *MSNBC*: [www.msnbc.msn.com](http://www.msnbc.msn.com)
- 11- *News Brief (European Media Monitor)*:  
<http://press.jrc.it/NewsBrief/worldedition/en/en.html>
- 12- *Newzbubble*: <http://www.newzbubble.com/index.php>
- 13- *News Explorer (European Media Monitor)*:  
<http://press.jrc.it/NewsExplorer/dayedition/en/latest.html>
- 14- *Newsmap*: <http://www.marumushi.com/apps/newsmap/newsmap.cfm>
- 15- *Notibits*: <http://www2.ing.puc.cl/~dcolle/notibits/index.htm>
- 16- *Panopticon*: <http://www.panopticon.ufba.br/>
- 17- *Portal Terra*: [www.terra.com.br](http://www.terra.com.br)
- 18- *TDC*: <http://www2.ing.puc.cl/~dcolle/notibits/index.htm>
- 19- *The New York Times*: [www.nytimes.com](http://www.nytimes.com)
- 20- *Topix.net*: [www.topix.net](http://www.topix.net)
- 21- *Urbi et Orbi*: [www.urbi.ubi.pt](http://www.urbi.ubi.pt)
- 22- *Washington Post*: [www.washingtonpost.com](http://www.washingtonpost.com)

## APÊNDICE II

## FORMULÁRIO DE OBSERVAÇÃO

## PARTE I

## IDENTIFICAÇÃO DO PRODUTO

1 - Nome \_\_\_\_\_

2 - URL:

3 - Tipo de produto:

- Produto do *mainstream* da mídia mundial
- Produto do *mainstream* da mídia nacional
- Produto de empresa não jornalística, mas de referência na Web
- Produto acadêmico/jornal laboratorial
- Produto experimental
- Blog*

4 – Especificidades que apresenta:

---

---

---

5 – Elementos que organizam a oferta informativa:

---

---

---

6 – Estrutura das peças informativas e dos recursos que incorpora:

---

---

---

7 – Há seções dedicadas à participação dos usuários como produtores de conteúdos?

Sim

Não

Em caso positivo, quais são as formas de participação e os nomes respectivos das seções?

---

---

---

**8** – Como é trabalhado o material de arquivo?

- Apenas em seção específica para pesquisa  
 Está agregado à estrutura das peças informativas, para aprofundamento e contextualização  
 Compõe novas seções da oferta informativa

Detalhamento:

---



---



---

**9** – Sobre a apresentação das informações:

- Em geral, segue padrões convencionais, relacionados à metáfora do impresso  
 Provê modos diferenciados de visualização para os conteúdos

---



---



---

**10** – Disponibiliza o conteúdo jornalístico para plataformas móveis?

- Sim  
 Não

Em caso positivo, para quais plataformas?

---



---



---

**11** – Categorias do Jornalismo Digital em Bases de Dados que apresenta:

- |                                      |                          |
|--------------------------------------|--------------------------|
| - Dinamicidade:                      | <input type="checkbox"/> |
| - Automatização                      | <input type="checkbox"/> |
| - Inter-relacionamento/Hiperlinkagem | <input type="checkbox"/> |
| - Flexibilidade                      | <input type="checkbox"/> |
| - Densidade Informativa              | <input type="checkbox"/> |
| - Diversidade Temática               | <input type="checkbox"/> |
| - Visualização                       | <input type="checkbox"/> |

**PARTE 2****1- DESCRIÇÃO:**

**- Dinamicidade**

**- Automatização**

**- Inter-relacionamento/Hiperlinkagem**

**- Flexibilidade**

**- Densidade Informativa**

**- Diversidade Temática**

**- Visualização**