

A recuperação de informações históricas e jornalísticas

A. Fernanda P. A. Wright
Alceu Antonio da Costa

RESUMO

O artigo apresenta metodologia de captação de informações, baseada no Projeto Leviathan desenvolvido pela Profa. A. Fernanda P. A. Wright junto ao Departamento de História/USP, para formação de um "Banco de Dados" contendo artigos de jornais já publicados.

Palavras chaves: informações jornalísticas, Projeto Leviathan, Banco de Dados.

RESUMEN

El artículo presenta la metodología de captación de informaciones, basada en el Proyecto Leviathan desarrollado por la Profesora A. Fernanda P. A. Wright junto al Departamento de Historia/USP, para la formación de un "Banco de Datos" conteniendo artículos de periódicos ya publicados.

Palabras centrales: informaciones periodísticas, Proyecto Leviathan, Banco de Datos.

ABSTRACT

The article presents methodology of information captation, based on Leviathan Project, developed by Prof. A. Fernanda P. A. Wright with the History Department/USP, to the formation of a "Datos Bank" containing newspapers articles already published.

Key words: journalistic informations, Leviathan Project, Datos Bank.

As novas tecnologias atualmente em uso nas redações dos jornais apressam facilidades incomensuráveis. Rapidamente é possível compor e redigir um texto para o jornal, o que na prática significa queimar muitas etapas.

É inegável que o uso do computador na redação traz vantagens, porém, apresenta inúmeros desafios. Assim, é que a informatização da redação do jornal acelera a produção do jornalista, mas introduz problemas sérios no tocante à parte informativa de caráter histórico.

Os "Bancos de Dados" atuais

Os principais jornais sempre mantêm seus arquivos contendo os números de jornais já publicados. Os jornais do século passado e começo deste, podem ser consultados em hemerotecas e arquivos públicos, mais destinados ao trabalho dos pesquisadores em ciências humanas e intelectuais em geral à procura da reconstituição de um passado relativamente distante. O problema para um repórter que tem que fazer uma matéria urgente pesquisando o arquivo tradicional, é que expende demorado tempo de pesquisa, o que em muitas das vezes inviabiliza seu artigo.

A expressão "Banco de Dados", comum em ciência da computação, é usada por muitas redações jornalísticas de maneira imprópria, pois leva a crer que se trate de massa de informações respeitável arquivada em fitas/discos de um Host para serem acessadas por terminais ou micros.

Na realidade, esses "Bancos" nada mais são do que uma tentativa canhestra de classificação de matéria por assuntos, formando pilhas intermináveis, facilmente desorganizadas durante o manuseio dos interessados.

A alimentação desses informes jornalísticos de um Mainframe ou Gateway pode ser tentada com vistas às informações armazenadas.

Poucos são os especialistas que dispõem de metodologia capaz de orientar de maneira uniforme a captação das informações desejadas. É lugar comum a afirmação de que o computador não "pensa", porém, desenvolve apenas o que lhe é oferecido.

Para tentar recuperar tais informações é possível usarem-se "pacotes de software", destinados ao processamento de textos. Neste ponto começa o problema da recuperação de informações históricas ou jornalísticas.

Armazenam-se as informações, porém, o roteiro para recuperá-las é tão complicado que a visualização da notícia, do artigo, tal qual foram redigidos inicialmente, é praticamente impossível de reproduzi-los.

Da Metodologia

Usando-se metodologia destinada a uniformizar a captação das informações, é possível compor um manual de procedimento bastante

simples capaz de servir de guia à tarefa fundamental da formação de "Bancos de Dados" realmente efetivos.

O Leviathan e suas lições

O projeto Leviathan foi desenvolvido pela Prof^a A. Fernanda P. A. Wright inicialmente junto ao Departamento de História/USP entre os anos de 1973 e 1978 para analisar a transição da influência econômica e política dos Estados Unidos contra a Inglaterra nos debates parlamentares brasileiros, entre 1828 à 1837. Começou afortunadamente pelo treinamento de um grupo de alunos de pós-graduação basicamente de Ciências Humanas, compreendendo arquitetura, geografia, história, jornalismo e, posteriormente, futuros engenheiros e matemáticos interessados. Os mesmos foram instruídos quanto à operação dos computadores, embora o objetivo fosse voltado para o uso da metodologia de captação das informações, o fator principal de êxito.

Pontos fundamentais do Leviathan

A leitura de um texto e a escolha dos *descritores* capazes de caracterizá-lo, constituiu o ponto central de todo o trabalho desenvolvido. A descoberta do *descriptor* foi feita após a compreensão de que o mesmo deveria partir da idéia do *homem em ação em seu meio ambiente*.

Foi assim que chegou à idéia de que um só *descriptor* nem sempre poderia refletir com precisão uma ação, sem envolver vários outros *descritores*. Isto porque a ação humana é mais das vezes "circular" e não "retilínea".

Ao traduzir idéias em palavras, ou radicais das palavras, com a álgebra booleana para associação das palavras ou parte do texto definitivo, é permitido ao coletor através de uma planilha, constituir uma verdadeira rede de subsistemas para armazenar e depois recuperar as idéias existentes por detrás das palavras, chegando mais perto da realidade histórica.

A necessidade do *descriptor* precisar englobar os diferentes domínios contidos num só pronunciamento, cria para o pesquisador a obrigação de distinguir os fatos "em si" dos seus "efeitos pertinentes". Isto é, não reproduzir o texto histórico pura e simplesmente transportando-o para a memória do computador e sim fazer com que o *descriptor* signifique o destino que se quer dar a ele como saída, ou insistir nos fins e não nos meios.

O conhecimento científico não emerge naturalmente da pesquisa bruta e sim constitui solução para palavras formuladas, as quais demandam um sistema coordenado de modo lógico.

Com os recursos materiais da época, usando computador de grande porte conseguiu-se no Leviathan formular uma planilha que captava cinco *descritores* para cercar uma atividade humana. Hoje, já temos computadores com mais velocidade, com maior capacidade de memória

possibilitando o uso de software mais refinados, como o usado no convênio USP/Scnado Federal para o projeto OSIRIS referente aos índices Onomásticos e Temáticos dos Anais do Parlamento brasileiro.

A captação das informações entre vários centros culturais a partir de um enfoque sistêmico é usada na Universidade de Michigan, desde 1963, para rcolher informações sobre o comportamento eleitoral e votações na casa do Congresso americano.

A grande revolução dos micros com a possibilidade de processamento e armazenamento de informações e a formação de redes viabiliza a aplicação da metodologia do Leviathan ao pesquisador e às instituições com poucos recursos financeiros.

As ações humanas nos pronunciamentos do congresso em 1828 à 1837 foram tentativamente incluídas em *descritores* de setorização a saber:

1. Administração Pública e Governo; 2. Comunicações; 3. Cultura; 4. Direito e Justiça; 5. Economia; 6. Educação; 7. Forças Armadas; 8. Meio Ambiente; 9. Obras; 10. Política; 11. População; 12. Previdência; 13. Relações Internacionais; 14. Religião; 15. Saúde; 16. Sociedade; 17. Trabalho; 18. Transporte.

Note-se desde já, que existem e existirão sempre acréscimos a serem feitos às subdivisões dos *descritores* apontados e existirão outros novos no futuro. Isto devido a progressos científicos nos softwares e hardwares ou nos novos computadores. Isto porque seus softwares exigindo menor grau de especialização dos usuários para manuseio aliam-se aos hardwarcos possibilitando acesso e tratamento das informações pelas redes digitalizadas computadas ou por pacotes a 1200 a 2 Mbits. Satélites e fibras ópticas para apenas apontar algumas conquistas mais rcentes, também, podem ser acionados.

Como encontrar os subdescritores de um texto

Para descobrirem-se os *subdescritores* de um texto basta aplicar a idéia geral comum às Ciências, ou seja, de que os fatos científicos, possuem *bases* e nele cntram *recursos e entidades* sendo que a manifestação final deve coincidir com o descritor desejado. Exemplo: no descritor *transporte* temos o conjunto "Base" onde estariam princípios, teorias e bases para o transporte.

Dentre os recursos a serem usados, teríamos pessoas ou rrecursos humanos. Nos recursos materiais, teríamos equipamento, patrimônio, recursos de hardware, tais como: máquinas ou entradas. Em entidades incluiríamos cousas como: agremiação, instituição, entidade patronal, trabalhista e sociedade, lembrando ainda, a possibilidade de existirem entidades ocasionais, tais como: comissão, congresso, seminários, exposição, feiras.

Dentro das manifestações estariam contidas caractcrísticas próprias da setorização dos descritos *transporte* que poderiam conter palavras como: aeroporto, ancoradouro, porto, estação, heliporto, terminal, base de lançamentos de foguetes enquanto os meios de

transporte, também do descritor *transporte* poderiam incluir avião, balsa, trem, animal, carroça, pedestre, automóvel, ônibus, metrô, embarcação.

As rotas e vias aí estariam incluídas como manifestação do descritor *transporte* podendo ser de vários tipos, tais como: via aérea, caminho, rodovia, ferrovia, fluvial, marítima, espacial, rua, alameda, com qualificadores a saber: taxa, preços, acidentes, depredações, paralisações, pouso, roubo, fiscalização, garantia, manutenção, proteção, frete, segurança, sinalização, baldeação, entroncamento, pedágio, entre outros.

A finalidade do *transporte* poderá ser considerado uma qualificação de todas estas manifestações do descritor geral de setorização.

Apenas para dar um último exemplo: esta finalidade poderá ser comercial, financeira, produção entre outros.

Aplicação da metodologia descrita ao jornal

Para orientação do pessoal encarregado da coleta de informações jornalísticas apresentamos aqui um sumário de como avaliar um jornal.

Isto leva em conta o recrutamento de pessoal não necessariamente ligado à profissão. Daí, a simplificação propositada das linhas gerais de um jornal.

Todas as outras instruções baseiam-se num sistema de informações que usem um "package" que à maneira do CICS/STAIRS da IBM, tenha condições de localizar palavras controladas e reconheça parágrafos numerados. Um manual de uso do sistema terá de ser desenvolvido a partir das instruções aqui contidas.

O jornal apresenta características específicas como um veículo de grande divulgação, influência e prestígio, pois suas idéias são lidas e assimiladas principalmente por uma determinada corrente de opinião formada à sua volta. Ele é caracterizado por três aspectos, quais sejam: periodicidade, atualidade e variedade. O jornal apresenta-se, ainda, como repositório de informações. Sendo um veículo escrito, produz mensagens destinadas a durar. Pode, por isso, ser considerado um documento destinado à reflexão e formula a mensagem complementada por imagens. Já quanto ao seu conteúdo, para fins de captação de informações poderíamos dividir o jornal em matéria editorial e ineditorial.

A matéria editorial não deve ser confundida com o artigo de fundo que é de responsabilidade da redação do jornal.

A matéria ineditorial pode, informalmente, constituir-se em matéria paga, ou seja, textos e ilustrações originadas fora da redação do jornal. Outros tipos de matérias não editoriais podem ser assinadas por colaboradores, porém, a redação sempre delas tem conhecimento.

Num primeiro momento, seria aconselhável captar a matéria editorial passando para outras, se houver decisão de fazê-lo e verbas suficientes para a empreitada.

A matéria editorial do jornal em geral pode ter três aspectos e ser: opinativa, informativa e interpretativa. Ela se apresenta ainda dividida em

diversos tipos, nos quais notaremos sempre a presença de um dos aspectos já citados, a saber:

- Editorial - É o mesmo que artigo de fundo ou principal matéria de um jornal, isto é, matéria opinativa que reflete o pensamento do editor e o ponto de vista do jornal sobre determinado assunto.
- Artigo - Comentário assinado no qual é importante a fundamentação podendo ter matéria informativa, interpretativa ou opinativa (ensaio).
- Crônica - Matéria assinada de caráter mais literário, leve, íntima, clara e precisa.
- Coluna - Sessão assinada e publicada com regularidade.
- Entrevista - Comentário ou opinião fornecida verbalmente a redação para publicação; é matéria opinativa.

Técnica de coleta

O formulário: o formulário utilizado na coleta das informações jornalísticas pode ser uma planilha no padrão utilizado no Projeto Leviathan, com algumas especificidades, que explicaremos a seguir, como exemplo:

I - Identificação: é registrada na linha 001, parágrafo 009, sendo obrigatório. São os seguintes os itens que o compõem:

- a) Responsável pelo texto: dependendo do tipo de matéria, o autor é o nome do jornal ou o jornalista, indicando-se por extenso. Se houver mais de um autor, indicar o nome de um deles seguido do símbolo &, e registrar os outros nomes no parágrafo "outros autores".
- b) Data do pronunciamento: dia, mês e ano da publicação do jornal, indicados por número e separados por barras.
- c) Origem: código do jornal, (ex.: "O Estado de São Paulo - ESP). Meio de Expressão: tipo de matéria a ser coletada devendo ser escrita por extenso, sem ser seguida de ponto.

- | | |
|---------------------|----------------|
| 1. Artigo | 7. Editorial |
| 2. Cartas à Redação | 8. Efemérides |
| 3. Crônica de Arte | 9. Entrevista |
| 4. Coluna | 10. Notícia |
| 5. Comunicado | 11. Reportagem |
| 6. Crônica | |

II - Descritor: Exemplo: 1. Descritor Geral
2. Descritor de Especificação 120, 140, etc.
Registrado no parágrafo 100, linha 002, e seguintes.

III - Manchete: Registrada no parágrafo 300. É o título que chama a atenção do leitor. Deve ser transcrito por extenso, seguida de ponto.

IV - Submanchete: Registrada no parágrafo 350.
Texto que se segue geralmente, abaixo da Manchete (com letras menores em relação a esta), citando aspectos essenciais da matéria. Deve ser transcrita por extenso seguida de ponto.

V - Síntese: Registrada no parágrafo 400.

Pode se iniciar com LEAD, ou seja, a parte da matéria geralmente colocada no início que responde às clássicas perguntas (nele subentendidas): Quem? Quando? Como? Onde? Por que?; e se for necessário, procede-se do mesmo modo que a síntese padrão do Projeto, principalmente quando o LEAD não estiver claro¹.

VI - Dados Adicionais: São registrados e subdivididos da seguinte forma:

Parágrafo 500 - Origem

Quanto a origem a matéria poderá ser:

- LOCAL - originada na cidade de publicação do jornal.
- NACIONAL - originada no restante do país.
- INTERNACIONAL - proveniente do exterior.

A seguir coloca-se a cidade, estado ou país da matéria em questão.

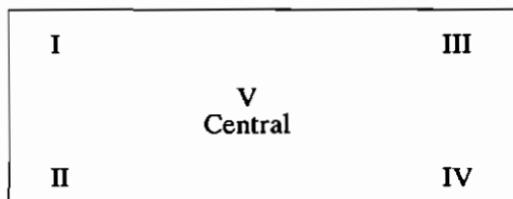
- DATA - só será colocada quando for data da redação.

Data da transmissão ou redação da matéria, se houver. Transcreve-se dia, mês e ano, separados por barra. Caso não apareça o dia, coloca-se o mês por extenso e o ano em algarismo. Exemplo: Novembro de 1930.

- AGÊNCIA - Transcreve-se a sigla da agência.

Parágrafo 510 - Descrição da Matéria:

Tamanho: a folha de um jornal é dividida em 5 quadrantes para fins de localização de matéria conforme figura abaixo.



A matéria ainda para efeito de coleta poderá ser:

- Pequena - menor que um quadrante.
 - Média - um quadrante.
 - Grande - maior que um quadrante.
 - Localização - a matéria poderá estar localizada nos quadrantes I, II, III, IV ou V.
 - Ilustração - poderá ser ilustrada com fotos, gráficos, etc.
- OBS.: Quando não houver ilustração não mencionar.
- Parágrafo 520 - Comentários:

Outros dados rápidos sobre a matéria. Quando for uma série de matérias, mencionar por ex.: terceiro artigo de uma série de artigos; Quando a matéria for resposta a outra, escrever RÉPLICA. Quando o

1 SF - Prefácio e Normas Técnicas de Organização e Apresentação do Trabalho em Índice Onomástico dos Anais da Câmara dos Deputados - 1826-1829 - Coedição Senado Federal - Ed. EDUSP, Brasília 1978 págs. IX à XIX.

autor se utilizar de um pseudônimo e o coletor estiver de posse de seu verdadeiro nome. Etc..

Os parágrafos que se seguem a este, possuem a mesma linha do formulário padrão. Seriam eles:

- Opinião de Outros: vide p. 9 dos Anais.
 - Tendências: vide p. 10 dos Anais.
 - Nomes Próprios: vide p. 12 dos Anais.
 - Outros Autores: vide p. 13 dos Anais.
- Registrado no
parágrafo
... linha ...

Referência bibliográfica. É anotada no parágrafo 900 e constitui-se:

- Título do jornal
- Local de publicação
- Data de publicação
- Número ou título do caderno, seção, suplemento
- Páginas e Colunas
- As informações seguintes são: registradas do parágrafo 100 à 800

100 a 499 - Descritores
500 - Manchete
600 - Submanchete
700 - Síntese

800 - Informações Complementares
900 - Bibliografia
910 - Descrição da matéria
920 - Comentários

O uso das novas tecnologias de comunicação e informação (NICI) abre um novo campo de tratamento da informação para os jornalistas e profissionais da comunicação.

A abundância de informação gerada pelos diversos meios torna cada vez mais difícil o acompanhamento dos fatos. Estes, hoje em dia, devido à complexidade da organização da sociedade, requerem dos profissionais uma interdisciplinaridade para seu tratamento, o que sem o auxílio da teleinformática torna cada vez mais difícil o seu exercício profissional.

Os pesquisadores mercadológicos tem no instrumental de tratamento da informação pela forma de acesso da mensagem através dos *descritores* exemplificados no Projeto Leviathan um grande aliado para sua análise.

São praticamente infinitas as aplicações da presente metodologia em outras áreas também dislumbradas no Projeto Leviathan. O campo pedagógico é privilegiado pelas concepções do Leviathan e os grandes projetos de informatização total das universidades, sem dúvida, seriam beneficiadas por essa concepção simples e efetiva da metodologia aqui exposta.