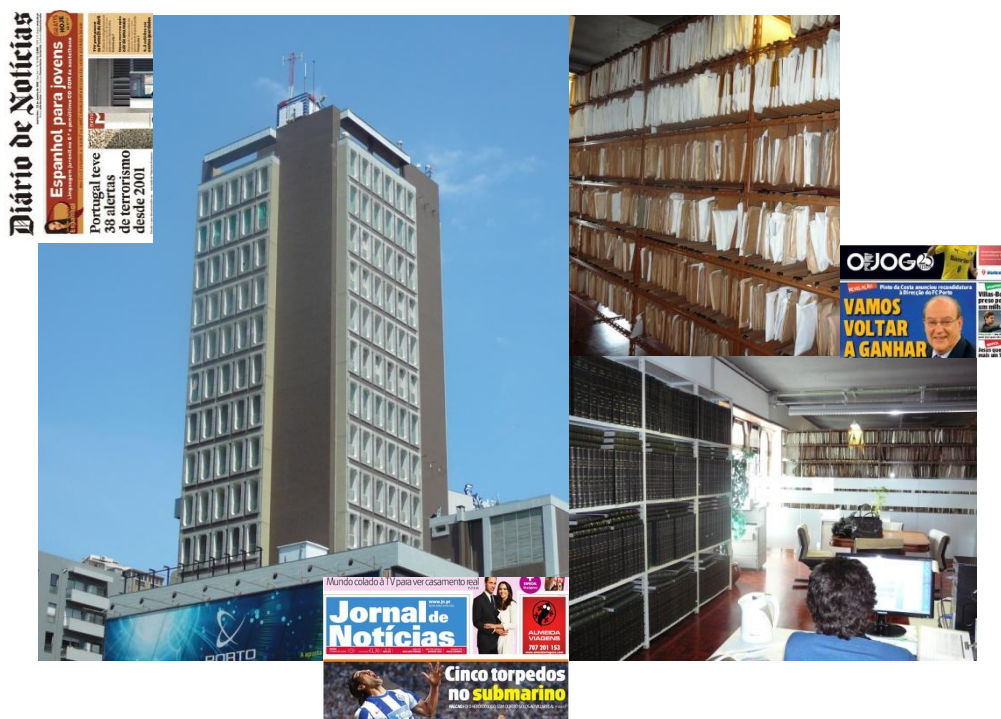


Maria José da Cruz Rodrigues

PRÁTICAS DE INDEXAÇÃO NA IMPRENSA

O Arquivo e Centro de Documentação do Grupo Controlinveste

Volume I



Universidade Fernando Pessoa

PORTO
2013

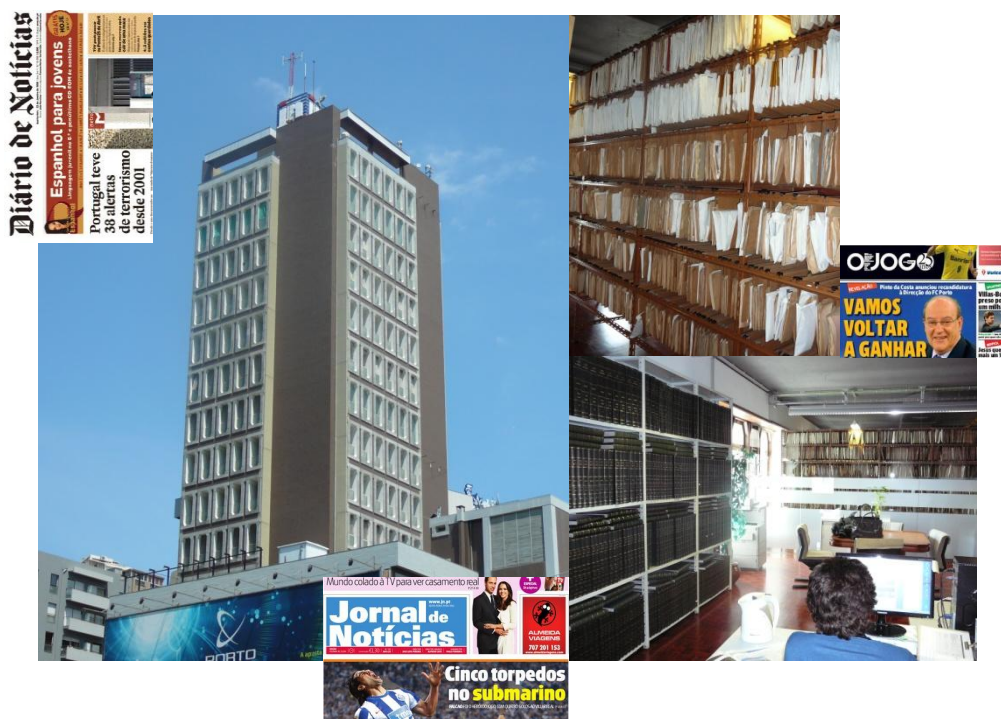
[em branco]

Maria José da Cruz Rodrigues

PRÁTICAS DE INDEXAÇÃO NA IMPRENSA

O Arquivo e Centro de Documentação do Grupo Controlinveste

Volume I



Universidade Fernando Pessoa

PORTO
2013

Maria José da Cruz Rodrigues

Práticas de Indexação na Imprensa

O Arquivo e Centro de Documentação do Grupo Controlinveste

Declaro que esta Dissertação é o resultado da minha investigação pessoal e independente. O seu conteúdo é original e todas as fontes consultadas estão devidamente mencionadas no texto e na bibliografia.

A candidata,

Porto, 23 de outubro de 2013

Dissertação de Mestrado
apresentada à Universidade
Fernando Pessoa, como parte dos
requisitos para obtenção do grau de
Mestre em Ciências da Informação e
Documentação.



RESUMO

A presente dissertação de Mestrado em Ciências da Informação e Documentação, intitulada “Práticas de Indexação na Imprensa” pretende mostrar de que forma a Ciência da Informação contribui na ação das Ciências da Comunicação no que diz respeito à utilização da informação, desde a sua recolha, tratamento e difusão até ao arquivo e recuperação. Neste sentido, é feita uma reflexão sobre o papel dos arquivos e centros de documentação dos meios de comunicação social em geral. Estes são responsáveis pelo tratamento documental dos conteúdos jornalísticos produzidos nas redações; não sendo possível dissociar esta função do conhecimento intrínseco do seu potencial informativo, que assenta, naturalmente, no valor de memória, narratividade e democratização da divulgação. Tudo é contextualizado no novo paradigma comunicacional que os meios de comunicação social percorrem nesta era digital em que vivemos.

Partindo do princípio de que a Ciência da Informação estuda a informação como um processo social, desde a génese ao armazenamento, a Indexação assume um papel preponderante desta prática social, visto que conduzirá à sua categorização/representação e recuperação, construindo a informação e transformando-a em conhecimento. Assim, é apresentado o *modus operandi* do Arquivo e Centro de Documentação do Grupo Controlinveste, no que concerne à política de indexação e recuperação de informação. Tal constitui o estudo de caso realizado com base na investigação *in loco*, cujos objetivos assentam no levantamento das variáveis inferidas nas linhas de atuação no momento da indexação e, conseqüentemente, na perceção das prioridades durante a recuperação da informação.

Para tal, foram analisados os diferentes tesouros construídos pelo Arquivo e Centro de Documentação do Grupo Controlinveste, tendo em conta as categorias temáticas que abrangem, assim como os termos, a estrutura e relações básicas. Após a análise de uma amostra aleatória de documentos deste arquivo, foi possível delinear as principais variáveis definidoras da política de indexação aplicada. Neste sentido, os mesmos conteúdos foram abordados quanto ao número de descritores por documento; à utilização



de Tesouros por documento; o recurso ao infraconceito e à harmonia do assunto em relação à secção do jornal.

Tratando-se de um aspeto peculiar deste estudo, foi dedicado uma especial atenção ao uso dos Infraconceitos, tendo sido feita uma breve reflexão e análise à sua relação na indexação de fotografias.

Em forma de súpula, foram registadas as linhas de atuação na indexação dos documentos analisados, aferindo da verosimilhança das declarações proferidas pelo Diretor do Arquivo e Centro de Documentação do Grupo Controlinveste na entrevista efetuada, que serviu de consolidação deste trabalho de investigação.

PALAVRAS-CHAVE: Ciência da Informação, Ciências da Comunicação, Imprensa, Política de Indexação, Arquivos dos Meios de Comunicação Social, Jornalismo.



ABSTRACT

The current Dissertation of Master's Degree in Information and Documentation Science, entitled "Practices of Indexing in the Press" is aimed to show how the Information Science contributes in the action of the Communication Science as regards the use of information, since its collecting, treatment and diffusion to its archive and consequent retrieval. In this sense, a reflection is done on the role of the archives and documentation centres of the media in general. These are responsible for the documental treatment of the journalistic contents produced in the newsrooms of those media, thus not rendering it possible to dissociate that function to the intrinsic knowledge of its informative potential, which relies in the value of memory, narrative style and democratization of the disclosure. Everything is put into context in the new communication paradigm the media wander in this digital era in which we live.

Assuming that the Information Science studies the information as a social process, from its genesis to its storage, the Indexing process takes on a ruling role of this social practice, since it will lead to its categorization/representation and retrieval, construing the information and transforming it into knowledge. Thus, the *modus operandi* of the Archive and Documentation Centre of the Group *Controlinveste* is presented, in what it concerns the politics of indexing and retrieval of information. Such constitutes the case study performed with the basis in the investigation *in loco*, whose aims are based in the collection of variables inferred in the lines of action in the moment of indexing and, consequently, in the perception of the priorities during the retrieval of information.

For that purpose, the different thesauri built by the Archive and Documentation Centre of the Group *Controlinveste* were analyzed, being taken into consideration the thematic categories they comprise, as well as the terms, its structure and basic relations. After the analysis of a small random demonstration of documents from this archive, it became possible to outline the main variables which define the politics of indexing applied. In this sense, the same contents were approached as to the number of descriptors by document; as to the use of Thesauri by document; the resource to *infraconcept* and the harmony of the matter to the newspaper section.



Being as it is a peculiar aspect of this study, a special attention was dedicated to the use of *Infraconcepts*; thus a brief reflection and analysis to its relation in the indexing of photographic images has been done.

In summary, the lines of action in the indexing of the analyzed documents were registered, going towards to the statements uttered by the Director of the Archive and Documentation Centre of the Group *Controlinveste* in the interview that was conducted, which served the purpose of consolidation of this research work.

KEYWORDS: Information Science, Communication Sciences, Press, Politics of Indexing, Media Archives, Journalism.



AGRADECIMENTOS

Um trabalho de investigação, como esta tese de mestrado, é fruto de um trabalho reflexivo, muitas vezes e necessariamente solitário. Porém, pouco se concretizaria sem a colaboração de algumas pessoas que, com a sua sabedoria e disponibilidade, ajudaram a construí-la. Estar disponível, nos dias de hoje, é uma virtude. Por isso, a todos que me acompanharam e ajudaram, deixo aqui a minha sincera gratidão.

Aos meus orientadores, Professora Doutora Judite Gonçalves de Freitas e Dr. António Regedor, pelo seu apoio e entusiasmo, deixando sempre uma palavra de confiança no meu trabalho, agradeço, por isso, as coordenadas fornecidas para encontrar o caminho desta tese.

Um profundo reconhecimento por toda a disponibilidade que, sem ela, não seria possível a realização deste trabalho, ao Arquivo e Centro de Documentação do Grupo Controlinveste, na pessoa do seu Diretor, Simões Dias, que sem qualquer constrangimento, abriu as portas deste arquivo, dando a conhecer toda a sua orgânica. Na verdade, esta nobre atitude demonstra o verdadeiro sentimento que um documentalista terá de desenvolver, o da partilha de experiências e conhecimentos. Acima de tudo, a minha sincera gratidão por tudo o que eu aprendi na observação efetuada nas instalações do Porto, sem dúvida uma dádiva para a vida.

O meu agradecimento sincero aos Professores que me acompanharam nestes dois anos de Mestrado de Ciência da Informação e da Documentação da Universidade Fernando Pessoa, que contribuíram para um maior enriquecimento na minha formação académica e pessoal.

Aos meus colegas de Mestrado em Ciência da Informação e Documentação, pela amizade demonstrada e, acima de tudo, por tudo o que com eles aprendi.

Por fim, porque os últimos são sempre os primeiros, à minha família, que sempre acreditou em mim e, por conseguinte, prestou todo o apoio para o equilíbrio necessário no desenvolvimento deste trabalho.



SUMÁRIO

Índice de Figuras

Figura 1 - A Informação para o Jornalismo e para a Ciência da Informação	7
Figura 2 – Modelo da pirâmide deitada	25
Figura 3 – Categorias dos termos de indexação	61
Figura 4 – Estudo de caso: esquema de trabalho	71
Figura 5 – Categorias dos termos de topo do Tesouro “Temas” do Grupo Controlinveste	87
Figura 6 – Áreas Temáticas do Tesouro “Temas” do Grupo Controlinveste	88
Figura 7 – Categorias temáticas do Microtesouro “Comunicação”	91

Índice de Ilustrações

Ilustração 1 – Edifício do Jornal de Notícias / Arquivo e Centro de Documentação do Grupo Controlinveste (Porto)	72
Ilustração 2 – Edifício do Diário de Notícias / Arquivo e Centro de Documentação do Grupo Controlinveste (Lisboa)	73
Ilustração 3 – Área de recortes do Arquivo e Centro de Documentação do Grupo Controlinveste (Porto)	73
Ilustração 4 – Arquivo e Centro de Documentação do Grupo Controlinveste (Porto)	74
Ilustração 5 – Menus das páginas em lha do “Jornal de Notícias” e “Diário de Notícias”	89
Ilustração 6 – Documento da imprensa: Exemplo 1	101
Ilustração 7 - Documento da imprensa: Exemplo 2	103
Ilustração 8 - Documento da imprensa: Exemplo 3	104
Ilustração 9 - Documento da imprensa: Exemplo 4	106
Ilustração 10 - Documento da imprensa: Exemplo 5	108
Ilustração 11 - Documento da imprensa: Exemplo 6	110
Ilustração 12 - Documento da imprensa: Exemplo 7	112
Ilustração 13 - Documento da imprensa: Exemplo 8	114
Ilustração 14 - Documento da imprensa: Exemplo 8	115

Índice de Gráficos

Gráfico 1 – Nº termos quanto à sua subordinação dos microtesauros “Actividade Económica”, “Comunicação” e “Trabalho”	96
Gráfico 2 – Nº de descritores por documento: exaustividade / especificidade	117
Gráfico 3 – Nº de Tesouros por documento	118
Gráfico 4 – Relação Descritores / Linguagem natural	120



Índice de Tabelas

Tabela 1 – Análise comparativa de diferentes linguagens documentais	56
Tabela 2 – Classes de Tesouros	59
Tabela 3 – Uso do qualificador e das notas explicativas	62
Tabela 4 – Abreviaturas dos termos num Tesouro	65
Tabela 5 – Análise comparativa das secções dos jornais com as Áreas temáticas do Tesouro Temas	89
Tabela 6 – Descritores e Não-descritores dos microtesouros “Actividade Económica”, “Comunicação” e “Trabalho”	92
Tabela 7 – Nº descritores com não-descritores dos microtesouros “Actividade Económica”, “Comunicação” e “Trabalho”	93
Tabela 8 – Natureza dos Não-descritores dos microtesouros “Actividade Económica”, “Comunicação” e “Trabalho”	94
Tabela 9 – Termos simples e termos compostos dos microtesouros “Actividade Económica”, “Comunicação” e “Trabalho”	95
Tabela 10 – Níveis hierárquicos dos termos dos microtesouros “Actividade Económica”, “Comunicação” e “Trabalho”	96
Tabela 11 – Relações hierárquicas genéricas e partitivas nos microtesouros “Actividade Económica”, “Comunicação” e “Trabalho”	97
Tabela 12 – Variáveis e objetivos da análise dos documentos da imprensa	117
Tabela 13 – Aplicação de infraconceitos em tipologias texto diferenciadas	118
Tabela 14 – Harmonia Assunto / Secção do jornal	119
Tabela 15 – Infraconceitos: desvios à NP 4036 (1992)	122
Tabela 16 – Categorias dos Infraconceitos.....	123
Tabela 17 – Exemplos de aplicação de infraconceitos em fotografias.....	125
RESUMO	I
ABSTRACT	III
AGRADECIMENTOS.....	V
INTRODUÇÃO	1
OBJETIVOS GERAIS	5
OBJETIVOS ESPECÍFICOS	5
I – CIÊNCIA DA INFORMAÇÃO: DIMENSÃO TEÓRICA E PRÁTICA.....	6
1. A Ciência da Informação e a sua relação com o Jornalismo	6
1.1- Informação, Conhecimento e Comunicação.....	7
2. O potencial informativo dos conteúdos jornalísticos da imprensa	10



2.1- Valor de memória	11
2.2- Narratividade	12
2.3- Democratização da divulgação	14
3. A imprensa na era digital: a mudança do paradigma comunicacional.....	16
3.1- A massa: um público sem consciência.....	16
3.2- A tomada de consciência das massas	19
3.3- O que se lê para além da nuvem	22
II – OS ARQUIVOS E CENTROS DE DOCUMENTAÇÃO DA IMPRENSA.....	27
1. O papel dos arquivos e centros de documentação dos meios de comunicação social.....	27
1.1- O profissional da ciência da informação e os conteúdos jornalísticos.....	28
2. Indexação: análise conceitual	33
2.1- Procedimentos e constrangimentos.....	33
2.2- Controlo da qualidade	37
2.3- Indexação manual, automática e semiautomática.....	40
3. Política de Indexação.....	46
4. Linguagem natural vs documental	49
4.1- Linguagem documental: livre ou controlada.....	51
4.2- Linguagem documental: pré-coordenada vs pós-coordenada	53
5. Tesouro: um instrumento de representação da informação	57
5.1- Classes de Tesouros	59
5.2- Classes de termos: composição e relevância	60
5.3- Estrutura e Relações básicas.....	63
5.4- Gestão da construção de um tesouro para conteúdos jornalísticos	65
III – ESTUDO DE CASO: PRÁTICAS DE INDEXAÇÃO DO ARQUIVO E CENTRO DE DOCUMENTAÇÃO DO GRUPO CONTROLINVESTE.....	71
1. Contextualização	72
1.1- Grupo Controlinveste.....	72
1.2- Arquivo e Centro de Documentação do Grupo Controlinveste: caracterização.....	73
1.2.1 - Arquivo físico	73
1.2.2- Circuito documental: jornal / arquivo	74
1.2.3- Arquivo digital e sistema de informação.....	75
2. Tesouros: gestão e classes.....	78
2.1- Gestão de construção	78



2.2- Classes e Tipologia	78
2.2.1- Tesouro AUTORES.....	79
2.2.2- Tesouro EMPRESAS	81
2.2.3- Tesouro EQUIPAS DESPORTIVAS	81
2.2.4- Tesouro FONTES.....	81
2.2.5- Tesouro GEOGRÁFICO	82
2.2.6- Tesouro INSTITUIÇÕES.....	83
2.2.7- Tesouro PERSONALIDADES	83
2.2.8- Tesouro SECÇÕES	84
2.2.8- Tesouro TIPO	84
2.2.9- Tesouro TEMAS.....	85
2.3- Categorias temáticas, termos, estrutura e relações básicas	90
2.3.1- Categorias temáticas no Microtesouro COMUNICAÇÃO	90
2.3.2- Termos, estrutura e relações básicas	92
3. Política de indexação dos conteúdos jornalísticos do arquivo digital do Grupo Controlinveste.....	98
3.1- Documentos da imprensa	98
3.1.1- Análise geral de 50 documentos	116
A- Número de descritores por documento	117
B- Utilização de Tesouros por documento	118
C- Recurso ao Infraconceito	118
D- Harmonia do assunto com a secção onde o documento está inserido	119
E- Relação dos descritores com a linguagem natural do documento	120
3.2- Fotografias e infraconceitos	120
3.3- Linhas de atuação na indexação dos documentos analisados.....	126
CONCLUSÃO.....	130
REFERÊNCIAS	134
1. Normas.....	134
2. Imprensa	134
3. Estudos.....	135



INTRODUÇÃO

A indexação, assim como nos é apresentada na NP 3715 (1989), é uma ação que consiste em descrever ou caracterizar um documento relativamente ao seu conteúdo, representando os seus assuntos numa linguagem documental. Com este processo, extraem-se os conceitos dos documentos através de uma análise intelectual sendo, de seguida, convertidos em termos de indexação. Portanto, trata-se de uma ação integrante do processo documental, na medida em que, através da análise de conteúdo, se faz, posteriormente, a recuperação da informação para o utilizador. Compreende-se, desta forma, que a indexação é um processo significativo para a comunicação e para o fluxo de informação em qualquer sistema de recuperação. No entanto, para que os assuntos integrados no acervo de qualquer organização sejam eficientemente recuperados, urge uma política de indexação devidamente estruturada e constituída no âmbito das necessidades do utilizador. Os objetivos dessa política são a definição das variáveis que afetam a indexação de documentos, assim como o estabelecimento dos princípios e critérios, que servirão de orientação ou, mesmo, de guia na tomada de decisões, no que diz respeito à escolha de descritores para a representação de um determinado assunto. Tal promove a racionalização do processo de indexação como, também, a consistência do mesmo e da pesquisa, no momento da recuperação de informação. Pode-se considerar que a política de indexação não é mais do que um tratado de procedimentos, explícito para o indexador e implícito para o utilizador. Essas decisões passam, entre outros fatores, pela escolha de uma linguagem documental que melhor poderá controlar os assuntos dos documentos de um acervo e, acima de tudo, que melhor representará a linguagem natural desses mesmos documentos, coincidente com a do utilizador no momento da pesquisa.

Esta linha tão ténue e complexa que existe entre a linguagem natural e a controlada é de extrema importância para o sucesso da representação dos assuntos de um acervo. Partindo do pressuposto que o indexador é o profissional que organiza a memória de um conjunto de documentos, a etapa da escolha de uma linguagem documental e seus princípios de atuação determina a qualidade da política de indexação e, consequentemente, a qualidade de um serviço documental.

Maria José Rodrigues

Se até aqui a palavra-chave foi *informação*, em contexto de tradução de assuntos e sua recuperação, impõe-se o cruzamento com o conceito de jornalismo ou comunicação social. Tendo em conta que jornalismo é a atividade de recolha, tratamento e difusão de informação através dos meios de comunicação social, interessa saber qual a política de indexação adotada num centro de documentação de conteúdos jornalísticos. De uma forma intrínseca, não é possível dissociar os interesses da política de indexação com o potencial informativo dos conteúdos jornalísticos. Estes apresentam características específicas que não devem ser esquecidas pelo indexador, para que a recuperação da informação encontre a dimensão social que envergam.

Entende-se os domínios da indexação e do jornalismo como áreas que convergem num instrumento comum: a informação. No entanto, também divergem na dinâmica que essa informação assume. Na comunicação social, a informação surge apontada a direções diversas, sendo o jornalista o profissional que a irá fazer veicular nos meios necessários até chegar ao cidadão comum. A informação, neste contexto, é difusa, espalha-se ao encontro das necessidades informativas da sociedade. Porém, o centro de documentação de um meio de comunicação social terá em mãos a difícil tarefa de controlar e classificar essa informação, através de uma tradução de assuntos de documentos que, mais tarde, poderão servir como base para mais uma difusão jornalística.

A este propósito é conveniente refletir na mudança de paradigma comunicacional na atual era digital. O facto da imprensa apresentar duplos formatos na divulgação dos seus conteúdos, em papel e em versão digital, é pertinente refletir o papel dos arquivos e centros de documentação dos meios de comunicação social no seu processo de tratamento documental, nomeadamente, na fase da indexação.

Após uma pesquisa bibliográfica, onde os conceitos de indexação e jornalismo não se cruzaram, apesar das suas afinidades, considerou-se pertinente investigar o *modus operandi* de um arquivo ou centro de documentação de um meio de comunicação social, no que concerne à política de indexação e recuperação de informação.

Mediante esta dimensão, pretende-se perceber de que forma a Ciência da Informação contribui na ação da Ciência da Comunicação. Relacionado com este objetivo, assenta o interesse em compreender o papel dos arquivos e centros de documentação dos meios de comunicação social. Para tal, é apresentada uma abordagem bibliográfica com o



intuito de cruzar esta ambivalência temática, de forma a preparar a apresentação de resultados concretos. Neste sentido, propõe-se entender a função das linguagens documentais no processo de indexação e recuperação da informação, assim como conhecer os objetivos e as variáveis da política de indexação de conteúdos jornalísticos de um arquivo dependente de um meio de comunicação social da imprensa. Para a consecução destes objetivos, será feita uma análise dos procedimentos e variáveis da política de indexação do arquivo digital do Grupo Controlinveste, a partir de uma metodologia baseada na observação direta e, com uma função complementar, de uma entrevista ao Diretor do referido arquivo e centro de documentação.

Por conseguinte, este trabalho estrutura-se em duas partes, sendo a primeira dedicada à revisão bibliográfica, onde se cruzam os conceitos-chave deste estudo: a indexação e o jornalismo. Nesta perspetiva, é apresentado um enquadramento teórico direcionado para a relação da Ciência da Comunicação com o Jornalismo, passando por uma reflexão sobre o potencial informativo dos conteúdos jornalísticos da imprensa. Claro está que a mudança de paradigma comunicacional que vivemos foi considerada, assim como o papel dos arquivos e centros de documentação dos meios de comunicação social. A estes aspetos junta-se uma análise conceitual da indexação e de fatores que lhe são inerentes, como é o caso da política de indexação; das linguagens documentais, como os Tesouros. Na segunda parte, é feita uma análise à política de indexação do Arquivo e Centro de Documentação do Grupo Controlinveste, no qual se insere jornais como o *Jornal de Notícias* e o *Diário de Notícias*, meios de comunicação social da imprensa escrita e de referência no nosso país. Este estudo, realizado em observação direta *in loco*, procurou encontrar a linha ténue e complexa entre a linguagem natural de um documento jornalístico da imprensa escrita e a linguagem controlada determinada no momento de indexação. Percebendo que existia uma política de indexação definida, cuja linguagem documental escolhida tinha sido o *Tesouro*, começou-se por fazer uma análise à sua gestão e classes, baseada nas diretrizes da NP 4036 (1992) sempre que se justificava, isto numa perspetiva de metodologia documental. Foi o *Tesouro* “Temas” que obteve uma maior atenção analítica, tendo sido feito um estudo às áreas temáticas mais abrangentes, áreas estas relacionadas com a estrutura intelectual das publicações. Deste tesouro, foram escolhidas três estruturas arborescentes, as quais foram apelidadas de microtesouros, tendo sido feito um levantamento dos termos preferenciais e não-preferenciais e, também, quanto à sua composição e relação de subordinação. Neste



seguimento, foram encontradas as variáveis da política de indexação do Arquivo e Centro de Documentação do Grupo Controlinveste, a partir da análise de variados documentos indexados no seu arquivo digital. Entre este, foi dada uma atenção especial às fotografias com o objetivo de aferir a relação desta tipologia documental com a utilização dos descritores *infraconceito*, aspeto tão peculiar da indexação de conteúdos jornalísticos.

Por fim, apresentam-se as conclusões deste estudo, complementando-se todos os dados recolhidos, tanto em observação direta como através da entrevista realizada ao Diretor destes serviço. Como é óbvio, com este estudo de caso não se pretende fazer generalizações, apenas consolidar e refutar teorias, lançar questões para futuras investigações ou, acima de tudo, acrescentar uma nova perspetiva em tudo o que concerne à política de indexação e respetiva recuperação de informação.



OBJETIVOS GERAIS

1-Mostrar de que forma a Ciência da Informação contribui na ação da Ciência da Comunicação.

2-Compreender o papel dos arquivos e centros de documentação dos meios de comunicação social.

3-Entender a função das linguagens documentais no processo de indexação e recuperação da informação.

4-Conhecer os objetivos e as variáveis da política de indexação dos conteúdos jornalísticos de um arquivo dependente de um meio de comunicação social da imprensa.

OBJETIVOS ESPECÍFICOS

1. Comprovar a relação de vizinhança entre Ciência da Informação e Ciência da Comunicação / Jornalismo.
2. Perceber a importância do trabalho do documentalista num arquivo e centro de documentação de um meio de comunicação social.
3. Apreender a gestão da construção dos Tesouros do Arquivo e Centro de Documentação do Grupo Controlinveste.
4. Conhecer a definição da política de indexação do Arquivo e Centro de Documentação do Grupo Controlinveste.

I – O VALOR DA CIÊNCIA DA INFORMAÇÃO NA ATIVIDADE JORNALÍSTICA: DIMENSÃO TEÓRICA E PRÁTICA DE UM HODIERNO RAMO DO CONHECIMENTO

1. A Ciência da Informação e a sua relação com o Jornalismo

Nesta abordagem introdutória sobre a relação de vizinhança que hoje pode estabelecer-se entre a Ciência da Informação e as Ciências da Comunicação em geral e o Jornalismo em particular, pretende-se realçar as principais características e explicar as contribuições da Ciência da Informação para a organização, classificação e recuperação dos conteúdos jornalísticos.

De acordo com Borko, um dos primeiros autores a definir o âmbito de atuação deste hodierno ramo do saber, a “Ciência da Informação é a disciplina que investiga as propriedades e o comportamento da informação, as forças que regem o fluxo informacional e os meios de processamento da informação para a optimização do acesso e uso. Está relacionada com um corpo de conhecimento que abrange origem, recolha, organização, armazenamento, recuperação, interpretação, transmissão, transformação e utilização da informação” [trad. nossa] (Borko, 1968: 3). De facto, nesta abordagem histórica e no que diz respeito à cientificidade da Ciência da Informação, absorvemos conceitos relacionados com o seu objeto de estudo que é a ‘Informação’, tais como fontes e recolha, organização e armazenamento, assim como recuperação e análise de conteúdo, terminando na fase comunicacional e seu uso. Também o mesmo autor indica que se trata de uma “ciência interdisciplinar derivada e relacionada com vários campos como a matemática, a lógica, a linguística, a psicologia, a tecnologia computacional, as operações de pesquisa, as artes gráficas, as comunicações, a biblioteconomia, a gestão e outros campos similares” [trad. nossa] (Borko, 1968: 5). Ora, neste alinhamento de ideias, compreendemos que o fluxo informacional, desde a origem até à sua divulgação, está naturalmente relacionado com o percurso operativo do Jornalismo. Considerando que este consiste numa atividade de recolha, tratamento / uso e difusão de informação, através de todo o tipo de meios de comunicação social, parece óbvio que a Ciência de Informação verte para o Jornalismo o seu *modus operandi*, garantindo-se como sua aliada. O Jornalismo recolhe a informação através de fontes específicas para depois fazer o seu tratamento com o intuito de a divulgar em linhas comunicacionais. Já a Ciência da Informação recolhe essa informação, que foi publicada ou divulgada, com o

objetivo de a tratar numa dinâmica organizacional, de forma a disponibilizá-la e recuperá-la.



Figura 1 - A Informação para o Jornalismo e para a Ciência da Informação

Na verdade, trata-se de duas atividades parceiras na medida em que a Ciência da Informação poderá desempenhar uma relevante função para o Jornalismo, assim como este será um alvo de estudo para aquela. E é nesta perspetiva que será feita uma abordagem comparativa do fluxo informacional de forma a melhor entender a referida complementaridade, considerando para o efeito a distinção conveniente entre informação, conhecimento e comunicação.

1.1- Informação, Conhecimento e Comunicação

Segundo Silva e Ribeiro, “a informação fica entre o conhecimento e a comunicação (...), envolve o processo cognitivo, as ideias, as impressões, os assuntos formalizados ou expressos por um código linguístico” (Silva e Ribeiro, 2002: 23). Neste sentido, a informação não existe sozinha, tem de estar contextualizada e organizada numa estrutura de assuntos que constitui o conhecimento. A comunicação, ainda para os mesmos autores, “apenas se actualiza quando emerge como processo de trocas entre os indivíduos”, sendo “a matéria dessas trocas” a informação. Portanto, “não há comunicação sem informação, mas pode haver informação sem comunicação, logo esta depende daquela” (Silva e Ribeiro, 2002: 26). Nesta perspetiva, entende-se que o Jornalismo necessita da informação devidamente contextualizada para poder comunicar de uma forma diacrónica e direcionada, este será o seu grande objetivo. Em contrapartida, a Ciência da Informação utiliza a informação, que pode ser denominada como ‘informação social’, com o intuito de a categorizar, contextualizar e organizar, a partir de diferentes suportes comunicacionais, podendo ser de uma forma assíncrona e multidirecionada.

Maria José Rodrigues

É evidente a distinção entre informação e comunicação, porém a sua fronteira com o conhecimento é mais ténue. Segundo Araújo, “a informação é uma prática social que envolve acções de atribuição e comunicação de sentido que, por sua vez, podem provocar transformações nas estruturas pois geram novos estados de conhecimento” (Araújo, 2002: 12). Eis aqui o cerne desta questão. Na verdade, o conhecimento evolui com uma nova informação. Esta poderá ser a consequência da interpretação de uma mensagem recebida, transformada numa nova informação no esquema comunicacional. Assim, ao Jornalismo importa recolher várias informações para que o acontecimento a divulgar se avolume, já que a informação, ainda segundo a mesma autora, é “um processo constantemente reconstruído pelo sujeito do conhecimento” (Araújo, 2002: 13). Assim, a informação não é um processo estagnado mas sim aberto, já que a realidade está sempre em mudança. Neste fluxo dinâmico, a Ciência da informação irá analisar o comportamento informacional pelos seus processos e propriedades, que são, segundo Silva e Ribeiro (2002: 39): “construção; comunicação e uso da informação”.

Ora, nesta análise parece que o Jornalismo e a Ciência da Informação processam em sentidos opostos, o que não corresponde à realidade. Apenas utilizam a informação em perspectivas diferentes. Contudo, visto que a segunda trata a informação como um fenómeno humano e social, o Jornalismo necessitará dela como uma ferramenta. Tendo em conta que o fenómeno informacional é um produto das relações sociais, o Jornalismo divulga esse mesmo produto que a Ciência da Informação analisa. Este processo acaba por ser útil para o Jornalismo na medida em que poderá recorrer à organização e à análise da informação para produzir o seu conhecimento, que consiste na narratividade de factos importantes para a sociedade que, num processo comunicacional, necessita que o seu recetor acredite e deposite credibilidade. Tal vai ao encontro da avaliação da informação a publicar que o jornalista terá de analisar.

Sendo assim, o Jornalismo materializa a informação em suportes físicos que constituem documentos gráficos, fontes de informação e conhecimento (jornais; revistas; registos de imagem e som; internet) através de uma interação social que é a comunicação. Quando essa informação entra no processo comunicativo e chega a um recetor que a interpreta, estamos perante a concretização do conhecimento. Tal vai ao encontro de umas das propriedades da Informação que é a transmissibilidade, ou seja, é potencialmente transmissível ou comunicável. A Ciência da Informação analisa,



conforme Silva e Ribeiro, “a determinação de constantes e de variações inscritas no fluxo informacional: como e porque se reproduz; qual a maior ou menor pregnância temática (potenciadora da pertinência); que ilações tirar de séries quantitativas de resultados; que tipos de perfil de utilizadores é possível determinar” (Silva e Ribeiro, 2002: 42).

Desta forma, entende-se que o objeto comum do campo de trabalho do jornalista e do documentalista é a informação. Ainda para Silva e Ribeiro, “o jornalista, enquanto indivíduo e enquanto membro de uma Organização – um jornal, uma rádio, um canal de televisão, etc. – é produtor / colector de informação e neste sentido é objecto de estudo da Ciência da Informação” (Silva e Ribeiro, 2002: 93).

Em sùmula, pode-se associar o Jornalismo ao conjunto de informações factuais comunicadas em massa. A Ciência da Informação capta esses factos com o intuito de as fixar numa categoria informacional. As dinâmicas são diferentes mas com interesses comuns.



2. O potencial informativo dos conteúdos jornalísticos da imprensa

Para abordar o conceito de conteúdo jornalístico, impõe-se, antes de mais, explicitar em que consiste o conteúdo editorial. Segundo Neveu, “designa o fluxo de textos ou de reportagens faladas ou filmadas que constitui um jornal (impresso, radiofónico ou televisivo). O conteúdo editorial, enquanto conteúdo, opõe-se a tudo o que é produzido fora da redacção (publicidade, comunicados)” (Neveu, 2005: 11). Desta forma, conclui-se que os conteúdos jornalísticos e editoriais são sinónimos, ou seja, são todos os textos produzidos numa redacção para serem publicados ou divulgados, cujo teor relaciona-se com factos ocorridos na atualidade do momento de escrita, isto na imprensa. Claro está que se trata de um suporte físico da informação no contexto comunicacional, que é expresso por um discurso que emite um sentido. Daí se chamar conteúdo a toda a produção informativa oriunda de qualquer meio de comunicação social, constituindo um documento gráfico e, conseqüentemente, uma fonte de informação e de conhecimento. Poderá ter diferentes suportes, sejam eles físicos ou digitais. No caso da imprensa escrita, é óbvio que são os signos linguísticos os portadores da informação, em diferentes tipologias textuais. A escrita é a principal expressão, embora a imagem auxilie a documentação da informação contida no texto. Essa imagem poderá ser uma fotografia ou gráfico ou mesmo um vídeo, este cada vez mais disponível nas plataformas digitais dos jornais e revistas.

Os conteúdos jornalísticos apresentam-se em suporte papel, no jornal ou revista, mas também em ambiente digital. A este propósito, o termo conteúdo tem sido bastante utilizado na discussão acesa entre os meios de comunicação social de todo o mundo e o motor de busca “Google”. É sabido que este disponibiliza, especificamente no “Google Notícias” ou “Google News”, vários títulos e respetivo conteúdo, provenientes de diferentes jornais, revistas e canais televisivos. Porém, esse agregador de conteúdos em ambiente digital não suporta qualquer custo nesta ação. Por esse motivo, já se começa a sentir algumas reações a esta situação, como aconteceu no Brasil em outubro de 2012, em que 154 jornais brasileiros anunciaram que pretendiam deixar de estar indexados ao Google News, uma vez que a empresa não quer pagar a utilização dos conteúdos, assim como foi noticiado no semanário Expresso, com o título “Jornais brasileiros impedem Google de usar notícias de borla” (Expresso, 20/10/2012). Ora, percebe-se, desta forma, que os conteúdos editoriais são a essência jornalística já que corresponde ao fruto do trabalho informativo dos jornalistas, sendo então a razão desta celeuma negocial. No dia

Maria José Rodrigues

07 de novembro de 2012, a plataforma digital do já referido semanário português anunciou a posição de Francisco Pinto Balsemão, presidente do conselho de administração da *Impresa* relativamente a este assunto, intitulado “Balsemão aguarda reação do Governo contra a Google” (Campos e Coelho, 2012, Expresso, 07/11/2012), tendo o mesmo explicado, na abertura da conferência “Media do Futuro”, decorrida em Lisboa, que o Governo alemão tinha anunciado que iria aprovar uma lei que obriga a Google a pagar aos *media* uma parte da publicidade pela utilização dos seus conteúdos. O mesmo é o que o *charmain* da *Impresa* (Expresso/SIC/Visão) pretende que o Governo português faça, afirmando “Não podemos continuar a diagnosticar os desafios estruturais e não lhes dar resposta. Os grandes agregadores de conteúdos não podem continuar a usar e abusar dos nossos conteúdos, a tirar-nos a publicidade e não serem obrigados a recompensar-nos” (Campos e Coelho, 2012, Expresso, 07/11/2012). Percebe-se aqui os interesses económicos envolvidos mas, acima de tudo, os interesses de autor dos jornalistas em todo o seu trabalho.

2.1- Valor de memória

Estes conteúdos lidam com os factos e com a sua divulgação. Importante será então falar do seu potencial informativo. Compreende-se que se trata de informação de grande realce para a história da Humanidade. Neles estão inscritos um valor de testemunho. A este propósito, Fernandes afirma que “o jornalismo não deve ser apenas uma atividade comercial, ou um diletantismo intelectual, deve antes possuir aquela dose indispensável de virtudes sociais que lhe permitem ter, na sociedade, um papel simultaneamente crítico e construtivo” (Fernandes, 2011: 47). De facto, o texto jornalístico representa uma fonte histórica e de informação. Para a História, é fundamental o que aconteceu, como aconteceu e, sobretudo, por que aconteceu. Sendo assim, o conteúdo jornalístico possui um valor histórico e permanente, preservando a memória de um lugar, de uma época ou de um acontecimento. É aqui que reside um dos seus potenciais informativos, ou seja, assente numa memória social e histórica que auxilia a compreensão do acontecimento descrito. Claro está que a forma como essa memória é percecionada pelo jornalista contextualiza-a num tempo. Este será um elemento agregador. Na verdade, para que um acontecimento seja descrito, terá de ser entendido pelo seu passado e pelo presente em que se situa, mesmo que esse passado seja muito recente. Nesta ordem de ideias, o conteúdo jornalístico assume uma perspetiva temporal que interfere na organização e apresentação do discurso jornalístico.

Maria José Rodrigues

Segundo Dalmonte, baseando-se em Santo Agostinho e Paul Ricoeur, a narratividade jornalística assenta no triplo presente. “O presente das coisas passadas refere-se à concepção histórica que (...) deve ser revisitada, pois é a historicidade dos factos que agrega sentido, atualizando o ocorrido. (...) O presente das coisas presentes é o facto enquanto tal; é o anúncio ou apresentação de um acontecimento. O presente das coisas futuras refere-se à influência do porvir que o acontecimento narrado pode fazer ressoar” (Dalmonte, 2010: 332). Claro está que um texto jornalístico acarreta uma dimensão temporal que abrange essencialmente o presente, mas para que tal seja devidamente interpretado, referenciais do passado terão de ser aplicados para, também, influenciar o futuro ou a sua continuidade, lançando o leitor numa expectativa constante. O papel destes conteúdos na difusão da memória é crucial, isto porque são eles que fazem com que ela se torne pública, isto é, acessível a todos, o que ajudará a que não seja esquecida. Sabe-se que a memória e a história sempre se confundiram e, com a influência dos meios de comunicação social, a história caminha em direção de um mundo constituído de memórias coletivas.

2.2- Narratividade

Os conteúdos jornalísticos comportam um outro potencial informativo que muito depende da forma como o acontecimento é exposto, ou seja, como ele é transferido para uma linguagem verbal. Tendo em conta a narratividade jornalística, Neveu *apud* Barthes quando aborda “as três dimensões da expressão” (Neveu *apud* Barthes, 2005: 83). A primeira é a “inventio” que consiste na recolha para escolher um ângulo, ou seja, o jornalista, antes de verbalizar o acontecimento, terá de recorrer a fontes e levantamento de informações adicionais, de forma a consolidar e contextualizar o descrito. Sendo a objetividade um marcador do discurso jornalístico, esta dimensão narrativa procura mostrar que os factos falam por si em detrimento da subjetividade de quem escreve. A segunda é a “dispositio”, que é a redação, o encadeamento dos factos, ideias e imagens. É aqui que entram as diferentes convenções da escrita jornalística, ou seja, os diferentes discursos que constituem formas narrativas, que veiculam uma visão diferente do mundo, que podem ser a notícia; a reportagem; a opinião; a entrevista; a biografia, entre outros. Por fim, a terceira é a “elocutio”, sendo esta explicada pelo estilo de quem escreve, pelas metáforas utilizadas ou pela utilização de frases feitas. Esta utilização singular do material linguístico terá de ser decifrado com referências do



passado residente na memória do leitor, o que poderá diminuir a expressão objetiva do conteúdo jornalístico.

Denote-se que a narratividade ou escrita jornalística é exercida por forças influenciadoras, como é o caso das condições e constrangimentos no momento da produção, assim como os objetivos do jornal tendo em conta o perfil do seu público-alvo. Certamente que diferente será escrever para uma faixa mais popular e com menos escolarização, que apenas pretende saber o essencial do acontecimento, ou para um público mais seletivo e exigente de um aprofundamento factual. Neveu refere que “apesar de fortemente codificada, a escrita de imprensa não é uma escrita estereotipada. A sua renovação é evidente, antes de mais, nas comparações feitas durante longos períodos de tempo. Estas mostram o processo de consolidação de uma dimensão analítica e interpretativa, para além de uma descrição positivista dos factos” (Neveu, 2005: 92). Na verdade, um dos textos jornalísticos que mais sofreu essa dimensão analítica ou interpretativa foi, sem dúvida, a reportagem, com a sua intelectualização, em oposição ao carácter meramente investigativo. O jornalista começa a acrescentar ao factual uma dimensão analítica e opinativa. Os artigos de opinião são um exemplo também dessa dimensão. Para Fernandes, “uma opinião que se revele verdadeira não deixa por isso de ter interesse, uma vez que só no confronto com outras opiniões se pode aprofundar o conhecimento” (Fernandes, 2011: 53). Na realidade, os artigos de opinião alimentam a memória dos conteúdos jornalísticos na medida em que apresentam diferentes perspetivas de um assunto.

A este propósito, interessa referir três modelos jornalísticos que marcaram a história do jornalismo. Vários estudos localizam na Grã-Bretanha e nos Estados Unidos da América a origem das práticas jornalísticas que hoje constituem norma de referência. Trata-se do modelo anglo-americano, como afirma Neveu, “a centralidade do factual (...), a predominância do discurso sobre a objetividade, construído à volta de uma transmissão dos factos, separando informação e comentário (...), fixando-se apenas nas respostas às questões “quem?; o quê?; como?; quando?; onde?”, são os parâmetros que o definem. Este modelo “procura uma descrição clínica dos acontecimentos, transforma os indivíduos e os factos em objectos de descrição fria e abstém-se de comentários (...), numa escrita sóbria e descritiva” (Neveu, 2005: 17). Outro modelo é o denominado “Jornalismo à francesa”, que se situa, como indica o mesmo autor, “entre literatura e

Maria José Rodrigues

política” (Neveu, 2005: 18). Na realidade, este modelo “fundamenta-se na mestria e brio do estilo(...), os conteúdos redactoriais, valorizando as críticas, os pequenos artigos e as crónicas, traduzem a importância do comentário, de um metadiscurso sobre a actualidade que privilegia a expressão de opiniões, transforma o acontecimento em pretexto para exercícios de estilo brilhantes e desenvoltos” (Neveu, 2005: 21). O terceiro modelo resulta de uma mescla dos dois primeiros, tratando-se do “New Journalism”. Neveu *apud* Tom Wolfe afirmando que teria de ser “um jornalismo que pudesse ser lido como um romance” (Neveu *apud* Tom Wolfe, 2005: 23). Ora, tal implica um trabalho quase etnográfico, pois pretende fornecer a descrição objetiva completa, ou seja, a mistura do factual com a vida emocional e subjetiva. De facto, trata-se de um modelo muito atual, já que a imprensa escrita mistura cada vez mais o lado factual com o emocional e, conseqüentemente, com o subjetivo. Claro está que todas estas perspetivas e modelos jornalísticos interferem como a realidade é transmitida e, necessariamente, com o registo da memória que qualquer conteúdo jornalístico potencia.

2.3- Democratização da divulgação

Neste momento, impõe-se a reflexão: será que a política editorial de um jornal reflete a realidade de forma a registar uma memória fidedigna do momento? Teoricamente considera-se que os conteúdos escolhidos para fazer parte de cada edição são uma visão da realidade. Contudo, serão mais uma construção da mesma. Fernandes refere que “no espaço público há pois que lidar com as muitas e diferentes escolhas que a cada momento estão a ser feitas por jornalistas. Dessas escolhas nem sempre resulta uma visão equilibrada ou mesmo aproximada da realidade. É frequente ocorrer uma espécie de «efeito de manada» e a maioria dos jornalistas começarem a copiar-se uns aos outros” (Fernandes, 2011: 11). Neste contexto, o potencial informativo dos conteúdos jornalísticos também se liga à democratização da sua divulgação. Ora, se os meios de comunicação social deverão funcionar como um contrapeso de forma a permitir que o povo se exprima ou, mesmo, que os vários poderes estejam presentes e se manifestem no mesmo espaço público, deverá a política editorial seguir interesses meramente económicos, cedendo a pressões de vária ordem? Para Pinto e Marinho, “o objectivo do jornalismo e dos jornalistas é o de prestar um serviço dirigido a cidadãos (e não consumidores), para quem é importante ter acesso a informação credível sobre acontecimentos relevantes sobre a vida pública” (Pinto e Marinho, 2003: 5). Desta



forma, os cidadãos estarão mais aptos a exercer funções de vigilância, fortalecendo a democracia. Assim como afirma Fernandes, o jornalismo também exerce a função de “*watchdog*”, isto é, “cão de guarda”, “alguém que está atento aos abusos do poder e às falhas no funcionamento das instituições”, fazendo funcionar “a máquina da opinião pública, permitindo que os cidadãos sigam os assuntos de interesse comum, avaliem os seus governantes(...)” (Fernandes, 2011: 49). Contudo, é notório que os conteúdos jornalísticos sofreram uma mudança significativa nos últimos anos. Percebe-se uma valorização da componente visual, tais como fotos e gráficos, assim como a compactação dos formatos e o aumento de rubricas *soft news*, ou seja, informações que não estão diretamente ligadas à atualidade, mas sim descrições, experiências de vida, informação prática em defesa do consumidor. A imprensa, numa atitude mais competitiva, alinha num modelo de telejornal, num ritmo alucinante de informação, principalmente no chamado *webjornalismo*, assente no suporte digital.

Com o aparecimento do jornal digital, a velocidade do ato de leitura mudou radicalmente. A imprensa escrita, para fazer face a esta tendência virtual, passou a ter menos texto, mais infografia e ilustrações, assim como inclusões publicitárias. Passou a haver necessidade de se ser menos analítico e mais explicativo, cuja informação deverá ser apresentada mais resumidamente e clara. Contudo, a complexidade de temas impede esta tendência. A imprensa vive na dicotomia brevidade/qualidade, o que nem sempre é fácil associá-las com êxito. Esta postura implica uma alteração no potencial informativo dos conteúdos jornalísticos, nomeadamente na sua escrita, na sua democratização de divulgação e na perspetiva temporal. A este propósito, Dalmonte refere que “a dimensão temporal é também um importante fator na organização do discurso que se desenvolve no ambiente web, como no *webjornalismo* ou jornalismo digital. A grande promessa gira em torno da possibilidade de oferta de notícias em tempo real” (Dalmonte, 2010: 340).

De facto, o jornalismo digital situa o jornalista, o acontecimento e o leitor num mesmo plano narrativo, o que também faz com que seja pertinente uma reflexão sobre a mudança de paradigma comunicacional que a imprensa atravessa nesta era digital.



3. A imprensa na era digital: a mudança do paradigma comunicacional

Vivemos uma fase de grandes mudanças no domínio comunicacional. A imprensa sobrevive à era digital numa postura de reinvenção permanente, após morte anunciada. Mas será que a comunicação de massas entrou totalmente em rutura? Parece que a imprensa ainda vive numa dualidade de dois paradigmas: o das massas e o da individualização. Adapta-se às novas tecnologias numa articulação com o modelo tradicional. Os dois suportes, o do papel e o digital, fundem-se numa tentativa de encontrar o equilíbrio, tendo em conta as preferências dos seus leitores. No entanto, a própria imprensa anuncia um vaticínio de extinção, como se constatou, no passado dia 25 de maio de 2012, no jornal “Diário de Notícias” com a publicação das declarações de Thomas Habe, presidente do grupo *Bertelsmann*, detentor do canal alemão RTL, onde refere que “O futuro da imprensa está na Internet paga”, tendo em conta a descida lenta e gradual das vendas de jornais. Esta previsão será realista ou apenas um alerta para a emergência de uma metamorfose?

Façamos então uma abordagem teórica do que foi e é a imprensa nas suas vertentes comunicacionais, para tentar compreender o posicionamento da mesma nos tempos em que vivemos.

3.1- A massa: um público sem consciência

Eis a denominação que Jean-Pierre Esquenazi atribui às massas, no seu livro “Sociologia dos públicos” e que serve como ponto de partida para a compreensão do conceito de sociedade de massa e da teoria hipodérmica da comunicação. Segundo Wolf, a massa é “uma estrutura social gerada pela evolução da sociedade capitalista, (...) é constituída por um conjunto homogéneo de indivíduos que, enquanto seus membros, são essencialmente iguais, indiferenciáveis, mesmo que provenham de ambientes diferentes” (Wolf, 1999: 9). Desta forma, podemos afirmar que a estrutura de massas é uma ilustração da sociedade pós Revolução Industrial e Pós-guerra, pois começou-se a pensar na produção em grandes quantidades sem ter atenção no que é diferente, singular ou individual. Ora, esta forma de organização social e intelectual também serviu os vários regimes totalitários, onde os valores individuais eram preteridos.

Maria José Rodrigues

A comunicação social, nomeadamente a imprensa, espelhou este formato comunicacional, acabando por adotá-lo. Percebemos que será a forma mais facilitada de transmitir informação, pois de uma massa sem consciência não se espera reação. A mensagem chega a um recetor formatado para assimilar, sem ter oportunidade e capacidade de refletir. E quando o pretendia fazer, o seu canal de comunicação não tinha eco de difusão. Os jornais viveram neste paradigma de capitalismo informacional, tanto durante a Galáxia Gutenberg, dominada pela escrita, essencialmente da imprensa, como na Galáxia Marconi, esta ocupada pelos meios eletrónicos, como o telefone, o cinema, a rádio e a televisão, mas que a imprensa sempre conseguiu acompanhar. Assim como afirma Serra, esta última galáxia comunicativa acentuou a massificação, na medida que passou a ter a capacidade de chegar, potencialmente, a toda a gente num contexto de mundialização, ilustrada pela metáfora da “aldeia global” e acessível a todos, de forma a acompanhar os acontecimentos e as “novidades” em tempo real (Serra, 2007: 62). Apesar da imprensa ter sentido alguma dificuldade em colmatar essa instantaneidade, acabou por fazer chegar a sua mensagem, apostando na qualidade do pormenor descritivo da sua narratividade. No entanto, as massas sem consciência eram as mesmas, isto é, sem poder de reação.

E foi assente nesta incapacidade de reflexão ativa que surgiram contestações a este paradigma dominante da massificação, como são exemplos disso a teoria crítica de Frankfurt; a teoria dos *media* da Escola de Toronto; os estudos culturais (*cultural studies*) da Escola de Birmingham; a semiótica de Roland Barthes e Umberto Eco; o estruturalismo de Michel Foucault, entre outros. Todas estas manifestações acabaram por focar a sua crítica, ou contributo, para uma outra perspetiva comunicacional, no conceito da manipulação que o controlo da informação massificada acaba por incutir. Por esse motivo, impõe-se abordar a força dos *media* e, mais precisamente, da imprensa sobre a opinião pública e os poderes políticos e económicos.

Sabemos que foi graças aos jornais que surgiu um público com a aceleração da circulação de opiniões pessoais. Apesar destas, inicialmente, terem tido uma movimentação difícil e controlada, a imprensa escrita foi um dos fatores que favoreceu o surgimento e desenvolvimento da opinião pública. Rieffel abordou o assunto controverso que constitui esse conceito. Para este autor, trata-se de uma “opinião partilhada por um grande número de indivíduos, uma opinião comum; trazida ao



conhecimento de todos e submetida ao juízo de todos, tornando-se assim pública. Para que haja opinião pública é necessária uma convergência de opiniões sobre um mesmo tema, uma expressão manifesta dos juízos e uma participação de um grande número de indivíduos” (Rieffel, 2003: 34). Percebendo este poder avassalador que tinha nas mãos, a imprensa criou condições para as massas dependerem dos jornais, construindo a opinião dominante, já que os *media* representam a principal fonte de referência para tudo o que sai da esfera pessoal dos indivíduos. Tal deixa transparecer que existe uma opinião dominante imposta pelos meios de comunicação social. Apesar do paradigma das massas ser um processo vertical e unidirecional, que procura uma mensagem igual para todos, estes ‘todos’ são iludidos num processo de influência. Para além disso, a imprensa escrita permite favorecer a circulação de ideias e de opiniões, dando a palavra a certas personalidades, incluindo nas suas colunas testemunhos de especialistas, que ajudam a construir essa mesma opinião pública, para além do exercício das sondagens. São os denominados ‘líderes de opinião’, muitos deles colaboradores dos meios de comunicação social, ligados hoje às Universidades, até pertencentes a outras áreas do conhecimento, que através do seu testemunho conferem cientificidade e credibilidade ao discurso jornalístico em formato de opinião. E serão todos estes fatores em articulação que exercem um poder nas massas.

A imprensa é considerada um poder, mais exatamente o ‘quarto poder’, por contraponto ao executivo, legislativo e judicial. Fernandes (2011) indica ainda que Ignacio Ramonet, diretor do *Le Monde Diplomatique*, tem vindo a defender que já não existem os outros três poderes, pois o primeiro estará a ser exercido pela economia, o segundo pelos *media* e o terceiro pela política. Fernandes discorda totalmente desta versão já que considera que a comunicação social não se articula “num centro decisório análogo ao executivo, ao legislativo e ao judicial, sendo condicionado e controlado por todos os outros” (Fernandes, 2011: 42). Em contrapartida, agora surge a ideia do ‘contrapoder’, o que é igualmente considerado, por este autor, inadequado e mesmo “instigador de comportamentos excessivos e desequilibrados, em regimes nos quais o poder executivo é exercido por delegação dos cidadãos e existe uma legitimidade democrática” (Fernandes, 2011: 43). No entanto, pondera-se sempre a reflexão de que a influência dos *media* coloca a questão do estatuto das democracias submetidas ao choque da mediatização. A propósito deste assunto, Ricardo Costa, como diretor adjunto de Informação da SIC Notícias, numa entrevista publicada pelo jornal Público de 3 de

Maria José Rodrigues

junho de 2006, afirmou que “os media têm muito poder, sobretudo a um nível imediato. (...) São estudados desde sempre como «manipuladores de opinião». Esse perigo ainda existe, mas hoje é muito mais premente o que decorre do imediatismo. É por isso que os media nunca se podem ver a si mesmos como «poder». Os *media* podem, quanto muito, ser um contrapoder. Mas devemos ter cuidado com o discurso ‘antimedia’ porque esconde quase sempre um desejo de controlar” (Público, 03/06/2006).

Afinal, quem controla quem? É a questão que se impõe na era do paradigma comunicacional da massificação, contrapondo uma notícia surgida no dia 05 de fevereiro de 2007, no jornal *Diário de Notícias*, cujo título é elucidativo “Governo de Lula critica poder dos media”, onde surge uma acusação à grande imprensa brasileira que “procura impor a sua vontade ao Governo e ao país, sem interesse nacional ou republicano, mas movida tão-somente por grupos económicos” (Mota, 2007, *Diário de Notícias*, 05/02/2007).

Resta assim concluir que neste paradigma massificado, o público sem consciência era formatado para uma opinião pública que se desenvolvia num panorama de poderes, fossem eles políticos, económicos ou comunicacionais.

3.2- A tomada de consciência das massas

Com a internet, as massas começaram a despertar da apatia, percebendo que poderiam ser ouvidas e participar na construção informacional. Outro fator importante foi a crescente escolarização e, conseqüentemente, uma maior capacidade de intervenção crítica. De facto, numa dinâmica de comunicação em rede, assistimos a uma fusão do interpessoal e da massa, cujos papéis dos intervenientes comunicativos vão sendo alterados. Os princípios da interatividade e da mediação começaram a vingar numa imprensa que estava habituada a selecionar a informação para as massas. O jornalista possuía esse filtro e começou a perdê-lo a partir do momento em que a informação passou a estar acessível a todos. Os jornalistas como *gatekeepers* perdem força, pois já não conseguem guardar a informação e escolher como e quando a divulgar. Ora, esta ‘abertura dos portões informativos’ estremeceu os *media*. A comunicação em rede obrigou a uma reinvenção do meio.

A imprensa portuguesa viveu, neste último quarto de século, mudanças significativas que a obrigaram a novas dinâmicas. Num primeiro momento surge a privatização dos

Maria José Rodrigues

órgãos de imprensa estatizados, o que lhe conferiu uma transferência de poderes, nomeadamente o económico. De seguida, o aparecimento das edições *online* dos principais jornais de referência, ou mesmo exclusivos nesta plataforma, como reação à invasão da internet. Nestes últimos anos, outro desafio lhes surge, que é o fenómeno dos jornais gratuitos, que não deixa de ser impulsionador de novas dinâmicas concorrenciais. É visível que o mercado da imprensa escrita se encontra em mutação e sob pressão.

Esta mudança de paradigma, que tem o seu cerne na democratização e inversão na hierarquia da comunicação, levou a imprensa a fazer uma *metareflexão* com a inclusão de secções dedicadas às transformações dos *media*, como é exemplo nos jornais Diário de Notícias e Público. Em 2005, no jornal Público, mais precisamente no dia 09 de outubro, surge uma notícia intitulada “Novos media estão a mudar os comportamentos colectivos” que alerta para a emergência de inéditos fenómenos de comunicação, como as smart mobs e a multiplicação exponencial de media móveis, que constituem uma nova paisagem e desafios, identificados como o telemóvel, a internet sem fios e definidos como “formas de acção coletiva organizadas a escalas antes impossíveis”. Para além disso, “as notícias também estão a ser alteradas pelas novas tecnologias. Os blogues estão lado a lado com os jornalistas na distribuição de informação na Internet, ajudados por serviços como o Google News, que seleciona notícias vindas de quaisquer páginas, jornalísticas ou pessoais, ou pelos serviços RSS (Really Simple Syndication)” (Cardoso, 2005, Público, 09/10/2005). E fica o apontamento de que a primeira imagem divulgada dos atentados de 7 de Julho em Londres não ter sido uma foto de um repórter, mas sim oriunda de um telemóvel e disponibilizada pelo Flickr, a mais popular aplicação de partilha de imagens do mundo.

Neste contexto, a imprensa viu-se impelida a aderir a estas tecnologias de cooperação ao assumir a vertente *online*. Num registo histórico, em julho de 1995, o Jornal de Notícias foi o primeiro jornal português a colocar a edição na *Web* e a começar a atualizar as notícias de última hora. O Diário Digital encetou a colocar a atualidade em linha desde julho de 1999. Em 2002 foi o primeiro jornal português a fornecer gratuitamente o alojamento de blogues. O Setúbal na Rede nasceu em janeiro de 1998 com o objetivo de levar a região de Setúbal ao mundo inteiro, tendo sido o primeiro jornal *online* português. O Jornal do Fundão foi o primeiro jornal português a cobrar pelos seus



conteúdos. Este semanário regional fechou o acesso livre em outubro de 2001. Enfim, nos anos 90, a rede estava instalada na imprensa escrita portuguesa, ainda antes das televisões e rádios. Contudo, numa fase inicial, a tendência foi para que os jornais *online* reproduzissem os conteúdos e formatos das suas edições em papel. Dez anos depois da primeira versão web de um jornal português, mais precisamente no dia 04 de junho de 2005, surge uma notícia no Diário de Notícias intitulada “Peritos dizem que futuro do jornalismo é online”. Tudo porque a internet é "um meio instantâneo indispensável" de acesso a notícias que "permite às pessoas ouvirem, verem e lerem o que querem, quando e onde desejarem" (Cardoso, 2005, Diário de Notícias, 04/06/2005). No entanto, salienta que o jornalista não soube acompanhar esta evolução, pois não adaptou o seu método de trabalho, já que não é possível trabalhar para *online* da mesma forma que para o papel, pois o jornalismo digital tem uma nova linguagem. E este aspeto constituiu um outro desafio para a imprensa digital.

Assistiu-se a uma segunda fase com a inclusão de conteúdos produzidos especificamente para a internet, tais como hiperligações para efeitos de aprofundamento da informação noutras páginas e *sites*, assim como a inclusão de registos sonoros e audiovisuais. Em termos de imediatismo, os jornais *online* começam a permitir uma atualização das notícias, já que incluem uma secção de ‘última hora’.

A interatividade também enriqueceu graças ao fenómeno dos blogues, que permitem a interação do público como, também, criar comunidades *online*. Isto, de facto, é o aspeto visível do novo paradigma da individualização comunicacional. Denota-se igualmente um crescendo da estratégia multimédia com o acesso a vídeos e galerias de fotografias.

Como se percebe, a aposta na interatividade e mediação entre a imprensa e o seu leitor é a grande característica da nova imprensa, seja no suporte papel ou digital. Vejamos o exemplo do Provedor de leitores, nomeadamente no “Diário de Notícias” ou mesmo no “Setúbal na Rede”. Esta figura reforçou a dimensão da cidadania dos leitores, tendo como principal missão atender as reclamações, dúvidas e sugestões, procedendo à análise regular do jornal, de forma a formular críticas e recomendações. Ora, nesta linha de ideias, Cardoso aponta quatro dimensões na mudança de paradigma comunicacional, sendo duas delas relacionadas com a individualização, que são as “novas dinâmicas de



acessibilidade de informação” e os “utilizadores como inovadores” (Cardoso, 2009: 36). Agora a imprensa obriga-se a ouvir o recetor que tomou consciência do seu mundo.

3.3- O que se lê para além da nuvem

O jornalismo digital proporciona a reciprocidade, já que o leitor tem a oportunidade de participar em fóruns de discussão, fazendo com que a informação se converta num intercâmbio. Na verdade, a comunicação jornal/leitor deixou de ser linear, no que concerne ao formato *online*. Num ambiente de hipertexto, torna-se difícil manter alguém muito tempo numa mesma página. Agora, o leitor decide como e onde encontra a informação, convertendo-se, mesmo, em cidadão participativo. A propósito deste conceito, surgiu um outro bastante discutível, o de “cidadão-jornalista”, e que foi posto em causa por Eduardo Cintra Torres, no jornal Público no dia 1 de novembro de 2005, com o artigo de opinião intitulado “Cidadãos quê?”. Isto porque não é possível confundir o cidadão que dá a sua opinião e alguma informação importante ou, até mesmo, que disponibiliza imagens, com a figura do profissional de comunicação. Tal não faz dele um jornalista, sublinha Eduardo Cintra Torres, “da mesma forma que, ao atender o telefone, o Presidente da República não é telefonista. Quem compra tábuas no Ikea e monta o móvel em casa não é marceneiro, quem faz uma transferência bancária numa ATM não é empregado bancário e quem enche o depósito de combustível em auto-serviço não é gasoleiro. Quer dizer, não tem essa profissão. O jornalismo é uma profissão.(...) A pessoa não se torna jornalista por participar no espaço público” (Torres, 2005, Público, 01/11/2005).

Daqui verte-se um outro conceito formado neste novo paradigma, o do *espaço público*. Este é constituído por todos os cidadãos que influenciam a formação da opinião através da sua participação. Assim como afirma Correia “o espaço público garante a relação entre a vida quotidiana e a vida política do Estado. Pelo menos uma parte das leis que regulam a atividade institucional do Estado é legitimada pelas práticas discursivas racionais dos seus destinatários e representantes e implicando mecanismos de formação de vontade e da opinião pública, que emergem do espaço público como instância autónoma de dinamização da sociedade civil” (Correia, 2006: 181). E assim percebemos que o espaço público é o lugar do debate e da argumentação, aberto à intervenção cívica, o que faz com que seja a partir dele que se construa a tão indefinida opinião pública, que deixou de ter raízes massificadas, adaptando-se ao novo paradigma comunicacional. Fernandes diz que “os cidadãos «entraram no círculo do poder» (...) ao



terem ganho o direito de livremente colocarem na rede aquilo que pensam” (Fernandes, 2011: 88). A aposta na interatividade é o ponto fulcral da nova imprensa, que reforça o papel do leitor no processo comunicacional. Numa postura vanguardista, o *Wisconsin State Journal*, um jornal dos Estados Unidos, em 2006, lançou um sistema de votação *online* para que os leitores pudessem escolher, entre um conjunto de artigos, qual deveria ir para a primeira página, procurando fomentar o interesse do público e responder à concorrência das notícias distribuídas através da internet. Já havia a consciência de que o avanço da tecnologia também permite aos leitores personalizarem a sua informação recorrendo a uma miríade de fontes. No entanto, levanta-se uma questão: será que este sistema interativo determina que o leitor esteja mais informado? Isto porque a seleção que faz poderá não ser a mais conveniente.

O que é certo é que a imprensa digital entrou no conceito da “nuvem”. Trata-se da rede de dispositivos que os utilizadores elegem para diferentes momentos do seu dia-a-dia, o que traz uma maior flexibilidade no acesso à informação. De facto, é um novo estilo de computação pessoal e que dá mais liberdade aos indivíduos, cada vez mais participativos, que numa multiplataforma consigam consultar e intervir quando e onde querem. A maioria dos jornais de referência já disponibilizam, em serviço pago, a edição digital em *ipad e Android* para *tablet e smartphone* ou em *ePaper (PDF)* para computadores pessoais. Ora, nesta situação, o leitor de um jornal poderá aceder, com os seus dispositivos tecnológicos em rede, aos conteúdos jornalísticos. Contudo, num sistema paralelo, é possível também encontrar a maioria das informações importantes do quotidiano em *sites* de acesso livre, daí começar-se a pensar que o papel do jornalismo como intermediário informativo começa a ser dispensável. Como afirma Fernandes, “estão a surgir novos centros de produção de informação de qualidade que, no limite, podem criar a ideia de que os jornais são redundantes e o jornalismo é descartável” (Fernandes, 2011: 91).

Desta forma, coloca-se a questão: mais do que os jornais, estará o jornalismo em crise? A propósito deste assunto, será pertinente refletir numa notícia publicada na revista “Notícias TV”, suplemento dos jornais Diário de Notícias e Jornal de Notícias, de 6 de julho de 2012, intitulada “Computadores substituem jornalistas”. Aqui é documentada uma nova plataforma informática norte americana denominada *Narrative Science*, que é capaz de escrever notícias a partir de uma programação com dados, estatísticas e gíria

Maria José Rodrigues

jornalística, tratando-se de notícias integralmente assinadas por computadores. Esta plataforma transforma dados em informação. Os assuntos deverão ser altamente estatísticos, o programa faz a leitura desses “mesmos dados, identifica cadências e exceções, analisa o significado e redige os textos finais” (Bernardino, 2012, Notícias TV, 06/07/2012). Segundo os criadores desta nova tecnologia, o objetivo não é exterminar o jornalista, apenas ajudá-lo. Pelo menos a revista *Forbes* já conta com algumas experiências nesta área, gerando algumas notícias automaticamente para a versão *online*.

Bem, assistimos a um esforço da imprensa em adaptar-se às novas tecnologias mas, ao mesmo tempo, também observamos leitores a dispensar o trabalho jornalístico pelo facilitismo do acesso informativo, assim como a própria imprensa a adotar um estilo mais mecânico e breve. Mas não será isto um fim consentido? O jornalista da imprensa deverá apostar na sua maior arma: a escrita. Esta é que poderá marcar a diferença. É certo que a imprensa viu-se desprovida do monopólio da escrita jornalística, já que as televisões e as rádios também o fazem nos seus domínios digitais. Para além disso, os leitores mais jovens parecem relegar para segundo plano os jornais, pois vivem numa era da imagem em movimento. Daí a necessidade do jornal em formato digital também disponibilizar vídeos e, acima de tudo, uma infografia atrativa e de qualidade. No entanto, terá de, paralelamente, investir na narratividade e não cair na tentação da escrita *fast-food*. Para além de escrever brevemente uma notícia, será profícuo recorrer a outras valências como selecionar, aprofundar, correlacionar, hierarquizar, comentar, interpretar, explicar, analisar, resumir e sintetizar, ou seja, o enriquecimento da informação. A este propósito surge o novo modelo de construção noticiosa proposto por João Canavilhas, onde a pirâmide invertida da escrita noticiosa deverá ser substituída pelo modelo da “pirâmide deitada”, argumentando que, desta forma, o utilizador tem a possibilidade de optar por diferentes sequências de leitura numa estratégia horizontal que vai de um espaço de menor informação para um de maior quantidade de informação, para cada um dos elementos estruturantes de uma notícia (o quê, quem, onde, quando, como e porquê). Assim, o leitor terá o poder de escolher a sua lógica de leitura de acordo com os seus interesses.

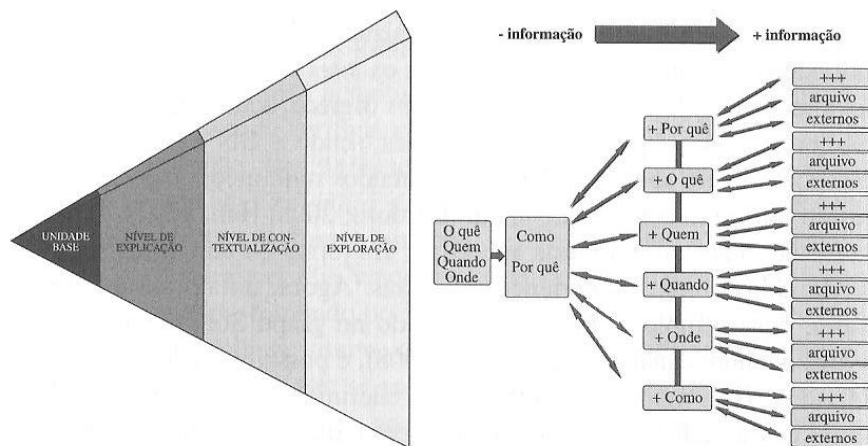


Figura 2 – Modelo da pirâmide deitada

E, desta forma, o jornalismo digital conquista a sua essência, deixando de assumir a sua faceta de fornecedor de informação para uma outra de fornecedor de conteúdos, numa dinâmica de funcionamento de agência noticiosa. Contudo, esta nova faceta da imprensa não poderá esquecer a credibilidade da informação que disponibiliza, passando pela transparência das suas fontes. Sempre foi uma preocupação da atividade jornalística, mas devido à velocidade informativa por uma busca da novidade permanente, elas foram muitas vezes descuradas nesta era digital. Recorrendo a uma rubrica do “Provedor de leitores” publicada no Diário de Notícias em 25 de maio de 2005, intitulado “As fontes anónimas do DN”, percebe-se, pelos estudos realizados na altura, que o anonimato crescia nos jornais de referência, sendo cada vez mais frequente expressões como “de acordo com as fontes contactadas...”, “fontes do gabinete do primeiro-ministro garantiram...”, “fonte de S. Bento negou...”, não confere uma clareza na construção e divulgação do facto informativo. Acrescenta que “raridade no recurso a fontes anónimas é um imperativo para manter a credibilidade nos órgãos da imprensa de referência. (...) Maior credibilidade será conseguida com as notícias, por regra, feitas com pessoas e instituições identificadas de forma precisa. Ou não será o jornalismo uma das raras formas de «inscrição» clara e frontalidade radical das sociedades democráticas?” (Abrantes, 2005, Diário de Notícias, 25/05/2005).

O jornalismo digital deverá apostar neste princípio da ética específica do jornalismo, como afirma Cardoso “da seleção e validação de factos” (Cardoso, 2009: 220). E a esta credibilização de conteúdos jornalísticos da era digital liga-se obrigatoriamente à função da imprensa como filtro da informação na internet, assim como foi sugerido na



Maria José Rodrigues

comunicação “O futuro incerto da internet: intercomunicar além do comércio e da publicidade” de Ricardo Jorge Pinto e Jorge Pedro Sousa, “essa poderá ser uma das funções futuras dos jornalistas: filtrar a informação na Net. Os seus órgãos de comunicação social poderiam ser as portas de entrada na Internet para quem está interessado em informação credível e útil. Mas, para isso, as empresas jornalísticas, além de disponibilizarem conteúdos, teriam de ofertar motores de busca onde os *links* apontados fossem apenas aqueles que contivessem informação efectivamente credível e útil” (Pinto e Sousa, 1998: 11). De facto, esta hipótese seria bastante viável, desde o momento que as empresas jornalísticas não condicionassem essa circulação de informação por motivos económicos, que regem o novo jornalismo, pois caso isso acontecesse, a manipulação da massificação regressaria. Lembremos que o novo paradigma comunicacional já está instalado ou, voltando a uma questão já colocada, será que a comunicação de massas entrou totalmente em rutura?



II – OS ARQUIVOS E CENTROS DE DOCUMENTAÇÃO DA IMPRENSA

1. O papel dos arquivos e centros de documentação dos meios de comunicação social

Nesta nova era digital que o jornalismo atravessa, os arquivos e centros de documentação dos meios de comunicação social desenvolvem um trabalho ainda mais específico que muitas vezes é esquecido. Frequentemente, eles constituem a base da produção noticiosa de um jornal, televisão ou rádio, fazendo parte integrante da pesquisa inerente à redação de um conteúdo jornalístico. Será, em muitas situações, uma das fatias da dimensão expressiva “inventio”, anteriormente referida, já que servirá de auxílio ao jornalista, que recorre ao arquivo ou centro de documentação para se fundamentar no desenvolvimento do acontecimento a noticiar. Nesta fase, o arquivo assume a função de memória viva, pois disponibiliza a informação, anteriormente publicada, de uma forma organizada.

A revista Notícias TV de 30 de março de 2012, publicação integrante dos jornais Diário de Notícias e Jornal de Notícias, publicou duas reportagens onde é ilustrada a importância do arquivo de um meio de comunicação social. Uma delas intitula-se “Como a RTP guarda a memória do país”, onde é descrito o funcionamento do arquivo da estação pública de televisão e todas as suas valências, documentando que “só em Lisboa, são necessários mais de três mil metros quadrados de depósitos climatizados para guardar toda esta memória, em que a História do País se confunde com a da televisão e da rádio nacionais” (Bernardino, 2012, Notícias TV, 30/03/2012). Por aqui, percebe-se a dimensão informativa que um arquivo desta natureza apresenta, cujo acervo é submetido a uma avaliação, seleção, tratamento documental, pesquisa e acesso, sendo o maior cliente a própria redação de Informação da RTP, registando-se que em 2011 “para fazer notícias e reportagens alargadas, os jornalistas da estação pública fizeram 17694 pedidos de imagens ao arquivo”, assim como “1624 pedidos enviados da área de programação e com destino (...) a entretenimento e documentários” (Bernardino, 2012, Notícias TV, 30/03/2012). Desta forma, percebe-se a função de bastidores que os centros de documentação desta natureza possuem no processo comunicativo da informação, com um público-alvo de consumo interno, ou seja, os jornalistas. No entanto, também serve um público externo, seja particular ou entidade coletiva, cuja compra de conteúdos comporta custos financeiros. A segunda reportagem

Maria José Rodrigues

tem como título “Depois do Adeus vai passar pelo DN”, onde documenta como a consulta ao arquivo do Diário de Notícias serviu para montar a cronologia da série televisiva de época “Depois do Adeus”, da RTP, centrada nos acontecimentos que ocorreram após o 25 de abril de 1974. Esse trabalho de pesquisa foi efetuado pela jornalista Helena Matos, que “recorreu ao arquivo do DN para pesquisar acontecimentos concretos”, como por exemplo “quando faltou a carne, o leite, quando houve greve nos transportes. Neste sentido, os jornais têm muitos pormenores do quotidiano”. A mesma ainda testemunha que “grandes arquivos não são só referentes aos jornais a que pertencem, mas são também arquivos da História de Portugal. É a vida de todos os dias do País, das pessoas... No caso do Diário de Notícias está tudo lá. E isso nenhum livro de História consegue dar” (Silveira, 2012, Notícias TV, 30/03/2012). Na verdade, esta última afirmação resume, de uma forma cabal, o papel dos arquivos e centros de documentação de um meio de comunicação social. Trata-se, de facto, de um serviço público, pois constitui um guardião da memória, mesmo que seja cada vez mais controlado por interesses financeiros.

1.1- O profissional da ciência da informação e os conteúdos jornalísticos

Torna-se agora obrigatório refletir nas funções do profissional da ciência da informação ou documentalista nestes arquivos. Tanto o arquivista, documentalista, bibliotecário ou gestor da informação fazem parte das profissões integrantes do campo de atividades da Informação, assim como a de jornalista. O documentalista é o profissional que gere a informação, regulando o seu fluxo pela organização. A informação que circula num arquivo de um meio de comunicação social é abundante, principalmente nesta era digital. A propósito da informação que circula em excesso nos dias de hoje, Pimenta considera que existe “uma boa oportunidade para o profissional da informação tornar o seu trabalho conhecido e imprescindível, já que gigantescos acervos de informação, sobre os mais variados temas – designados pelo nome genérico de conteúdos – circulam hoje, em escala planetária e de forma acelerada, por meio da Internet e das novas *mídias* eletrónicas” (Pimenta, 2002: 137). De facto, o trabalho quase invisível destes profissionais ocupa um lugar primordial na construção do conhecimento e sua acessibilidade, constituindo uma valorização dos conteúdos jornalísticos.

Desta forma e a partir das várias facetas do potencial informativo dos conteúdos editoriais, anteriormente referidas, urge analisar o papel do documentalista no

Maria José Rodrigues

tratamento da informação dos conteúdos jornalísticos para constituição de um arquivo devidamente organizado e recuperável. Este trabalho obedece necessariamente a normas de descrição arquivística, que não serão aqui explanadas. O que interessa é compreender o cuidado que os profissionais da informação deverão ter durante esse processo (avaliação, descrição, indexação e divulgação) pois são conteúdos informacionais que emergem de um processo comunicacional.

Sendo assim, ao documentalista, no momento da análise do conteúdo jornalístico para efeitos de arquivo, deverá reter a narratividade em que fluem essas informações. Se se tratar de um texto num modelo anglo-americano, cujos factos são normalizados numa escrita objetiva e utilitária, a Ciência da Informação absorve a sua informação de uma forma mais mecânica e igualmente objetiva. No entanto, os conteúdos jornalísticos não se cingem apenas a essa recolha de factos. Cada vez mais, o ‘jornalismo à francesa’ aplica o engenho e arte na narratividade, o que faz com que a informação a tratar para a Ciência da Informação necessite de uma mestria mais contundente. Encontrar o facto e o seu assunto em conteúdos que valorizam a crítica e a opinião exige um trabalho mais refinado dos profissionais da Ciência da Informação. No fundo, a construção do produto final do conteúdo jornalístico também é essencial para o documentalista, pois será através dele que a informação irá ser organizada, tendo de decifrar os propósitos do jornalista no momento da escrita. Na verdade, o profissional da Ciência da Informação terá de possuir sensibilidade estética para não confundir facto com opinião, podendo mesmo, numa fase de análise de conteúdo, eleger assuntos menos importantes ou, mesmo, desconsiderando outros essenciais.

Tratar informação ou textos de jornais diários, semanários ou mesmo de revistas exige pré-requisitos diferentes. Segundo Neveu, “o jornalismo das revistas é (...) terreno de eleição das *soft news*. (...) A maioria das revistas pode definir os seus conteúdos editoriais de uma forma relativamente desligada de uma actualidade feita de acontecimentos imprevisíveis” (Neveu, 2005: 39). O jornalista das revistas deve ir ao encontro dos interesses dos leitores de uma forma mais sensível. Assim, neste prisma, o profissional da informação terá de saber interpretar os objetivos editoriais de determinado conteúdo. Para além disso, tem de estar consciente das mudanças inerentes às políticas editoriais, pois será ele que deverá saber eleger os factos primordiais a classificar, organizar e arquivar numa dimensão de memória e difusão.



O documentalista consegue traçar um perfil do campo ou campos de abordagem que um jornal se dedica ou dá mais atenção. Naturalmente, consegue dar maior ênfase no tratamento da informação a temas também eles mais destacados nos conteúdos editoriais. Existem diversos fatores que influenciam esses campos de destaques, como os políticos, publicitários ou audiométricos. Como saber quais os assuntos a descrever com mais pormenor? Certamente aqueles aos quais foi dada maior relevância na publicação. Serve de exemplo a atenção superior que é dada ao futebol, isto no panorama desportivo, em detrimento das restantes modalidades.

Neste contexto, os conteúdos jornalísticos ou editoriais têm uma dimensão informacional relativa aos interesses do público ou, mesmo, financeiros. Neveu aborda o “peso da legitimidade dos próprios serviços” associado à “sociologia dos leitores”, afirmando mesmo que “todo o jornalista está estabelecido numa estrutura ramificada empresa-redacção-rubrica” (Neveu, 2005: 48-49). A hierarquia das secções varia em função dos contextos sociais e das publicações. Na verdade, o profissional da Ciência da Informação tem de entender a estrutura intelectual do meio de comunicação social que serve durante o seu trabalho. Este produz informação em conteúdos jornalísticos com objetivos diversificados, sejam eles económicos, políticos ou somente factuais. A imprensa escrita vive um jornalismo de mercado, cuja função do jornalista não se limita à transmissão de informação. Sendo assim, impõe-se pensar que o documentalista obriga-se a conhecer a política editorial da publicação. Caso não o faça, o tratamento da informação será desfasado das reais necessidades dos utilizadores que são os jornalistas dessa mesma redacção.

Como já foi referido, nem sempre a divulgação dos conteúdos editoriais transparecem a realidade na sua totalidade. Então, deverá o documentalista preocupar-se com essa transparência da realidade exposta nos conteúdos? A sua função será sempre tratá-los para disponibilizar num catálogo. Mas estará a constituir uma pseudomemória? Contudo, não poderá nem deverá colocar-se numa postura avaliativa do trabalho do seu colega de informação. Não é essa a sua função. Então, coloca-se a questão: o trabalho de um documentalista é democrático? Daqui poderíamos falar, por exemplo, de como indexar as notícias censuradas em tempos ditatoriais. Sabendo dos cortes a que estavam sujeitas, como tratar essa informação? Filtrar apenas os assuntos presentes ou ir mais além?

Maria José Rodrigues

Percebe-se que o jornalismo é entendido enquanto serviço público de qualidade. O mesmo deverá ser encarado nos arquivos e centros de documentação dependentes dos meios de comunicação social. Trata-se de uma responsabilidade social que deverá ser assumida, mesmo que de instituições com interesses privados se trate. Compreende-se que o imediato é pensar que tem de haver lucro, os serviços terão de ser rentáveis, encaminhando-se para a venda de conteúdos do próprio arquivo. No entanto, se a gestão não for feita num sentido de serviço público e responsabilidade social, certamente que esses mesmos conteúdos não serão vendidos. Os dois objetivos têm de ser aplicados simultaneamente. Investir na qualidade é sinónimo de credibilidade. Segundo Pinto e Marinho, na relação entre jornalismo e a sociedade, “a qualidade do jornalismo constitui, a um tempo, terreno de expressão e factor de qualidade de vida individual e colectiva, nas sociedades democráticas” (Pinto e Marinho, 2003: 7).

Se considerarmos a imprensa o quarto poder, então o documentalista tem em mãos uma responsabilidade maior no que diz respeito ao tratamento da informação como elemento da memória. Este profissional assume o papel de guardião da memória, daí a importância do tratamento da informação dos conteúdos jornalísticos. No entanto, partindo do pressuposto que um arquivo ou centro de documentação deve servir primeiramente a administração e só depois a História, estas instituições terão de encontrar o ponto de equilíbrio entre a produção informativa e os seus interesses e o legado para a memória como valor histórico que os conteúdos jornalísticos encerram. É urgente que os Arquivos e Centros de documentação dos meios de comunicação social trabalhem com um intuito de serviço público. Humanizar serviços, valorizar a memória social e humana irá servir de argumento a qualquer interesse financeiro. De facto, vender conteúdos e deles fazer negócio é valorativo, pois invoca interesse para qualquer entidade comunicacional. Mas não esqueçamos, e os dias de hoje ensinam-nos, que sem impulso afetivo, as instituições não funcionam e não serão valorizadas. É mais fácil cortar números do que eliminar pessoas. Quando os conteúdos jornalísticos passarem a ser meros números, toda a sua dimensão humana de que se servem será esquecida. E, então, a memória se reduzirá a meros centimos sem qualquer valor, apenas servindo a lógica de mercado.

A este propósito, parece pertinente citar a frase do poeta, periodista e ensaísta polaco Ryszard Kapuscinski: “Quando se descobriu que a informação é um negócio, a verdade



Maria José Rodrigues

deixou de ser importante”. Urge refletir nesta mudança a que assistimos nos meios de comunicação social, que cada vez mais se regem por um mercantilismo jornalístico que, por inerência, alastrar-se-á à funcionalidade dos seus arquivos e centros de documentação.

2. Indexação: análise conceitual

Como foi anteriormente referido, a Ciência da Informação estuda a informação como um processo social, desde a génese ao armazenamento. Neste contexto, a indexação assume um papel preponderante desta prática social, visto que conduzirá à sua categorização e, conseqüentemente, recuperação, construindo a informação e a sua transformação em conhecimento.

A indexação, segundo a Norma Portuguesa 3715 (1989), que tem como objetivo normalizar o método para a análise de documentos, determinar o seu conteúdo e seleccionar os respetivos termos de indexação, é apresentada como uma ação que consiste em descrever ou caracterizar um documento relativamente ao seu conteúdo, representando os seus assuntos numa linguagem documental. Com este processo, extraem-se os conceitos dos documentos através de uma análise intelectual sendo, de seguida, convertidos em termos de indexação. Portanto, trata-se de uma ação integrante do processo documental, na medida em que, através da análise de conteúdo, faz-se posteriormente a recuperação da informação para o utilizador. Conforme Lucas, “o processo de indexação é aquele que identifica o assunto de que trata o recurso, utilizando para isso, termos extraídos de um vocabulário controlado para descrever e representar o seu conteúdo temático”. Realça ainda que “é o ato de identificar e descrever o conteúdo de um documento ou recurso com termos representativos dos seus assuntos e que constituem uma linguagem de indexação” (Lucas 1996: 86).

Na verdade, a indexação é um processo de organização do conhecimento, ou seja, tem um carácter intelectual que envolve atividades cognitivas na compreensão de um documento e sua representação. Importante será também afirmar que a indexação é parte integrante de um dos momentos da cadeia documental, o do tratamento técnico documental, mais precisamente a descrição intelectual de um documento.

2.1- Procedimentos e constrangimentos

Ainda segundo a Norma Portuguesa 3715 (1989), a indexação divide-se essencialmente em três fases e que são:



1- Análise do documento e definição do seu conteúdo

Em primeiro lugar, o indexador deverá identificar os conceitos que melhor representam o conteúdo do documento, de uma forma precisa para não produzir ruído, isto é, todos os conceitos eleitos deverão ser relevantes no momento da pesquisa. É claro que este processo não implica, necessariamente, a leitura integral do documento, mas sim algumas das suas partes essenciais, tais como o título, resumo, sumário, introdução, conclusão, legendas, ilustrações ou diagramas. Isto nos documentos escritos, pois os não escritos, como é o caso de documentos audiovisuais, visuais, sonoros ou objetos, terão procedimentos diferentes. A este propósito, os conteúdos jornalísticos compreendem estas duas componentes, o que lhes confere alguma complexidade e um cuidado mais exigente no momento da indexação.

2- Identificação e seleção dos conceitos representativos do conteúdo

Nesta segunda fase da indexação, como está explícito na NP 3715, “os organismos devem construir grelhas de identificação que contenham os critérios considerados importantes na área abrangida pela indexação” (NP3715, 1989: 5). Desta forma, compreende-se que o ato de indexar um documento, seja ele textual ou não, interfere na recuperação do seu conteúdo. De facto, a indexação não é um ato isolado no tratamento documental, pois qualquer decisão neste momento implicará a representação da informação desse mesmo documento. O utilizador assume o papel principal neste processo intelectual. Ainda segundo a mesma norma, o indexador não tem necessariamente de considerar todos os conceitos identificados ao analisar o documento. A escolha desses conceitos depende do fim para o qual os termos de indexação vão ser utilizados. Ora, este fim depende dos interesses do destinatário e das possíveis questões que podem ser feitas ao sistema de informação. É aqui que estão assentes os princípios da exaustividade e da especificidade. Segundo a já referida norma “a exaustividade está ligada ao número de noções que foram consideradas e que caracterizam o conteúdo integral dos documentos”. Já a especificidade “está ligada à exactidão com que um determinado documento é representado por um termo de indexação” (NP3715, 1989: 6-7). Importante será saber quais são os interesses



do público-alvo, se apenas desejam recuperar os conteúdos dos documentos com assuntos específicos de forma a precisar a recuperação da informação pretendida ou, então, os vários assuntos que dele façam parte, de forma a recuperar mais exaustivamente o que procura. No que diz respeito aos conteúdos jornalísticos, o organismo deverá definir se interessa indexar uma notícia ou reportagem exaustivamente, atribuindo os vários assuntos que neles estão contidos. Por outro lado, poderá ser mais profícuo indexar pelo conceito para que a recuperação da informação seja mais exata, refletindo a objetividade deste tipo de textos.

3- Escolha dos termos de indexação

Esta última fase da indexação corresponde à tradução dos assuntos, eleitos nas fases anteriores, através de um instrumento de indexação que são as linguagens documentais. Estas poderão ser índices; listas alfabéticas de assuntos; listas de cabeçalhos; tesouros e classificações. Estes instrumentos auxiliam o indexador no momento da escolha dos termos que irão representar o assunto do documento, servindo também de ferramenta para o utilizador no momento da pesquisa. Na verdade, trata-se do elemento comum e integrador das ações do indexador e do utilizador. Contudo, tanto a utilização como a construção destes instrumentos requerem determinadas regras que se relacionam com a permanência dos termos já integrados e com a inclusão de novos a partir de conceitos recentes, o que lhes concede uma característica dinâmica e pouco estanque. Num arquivo ou centro de documentação de um meio de comunicação social e tendo em conta que se trata de conteúdos com um potencial informativo tão específico, será pertinente a construção dessas linguagens, já que os utilizadores são essencialmente os jornalistas, o que exige uma aproximação à linguagem jornalística.

Percebe-se que os dois primeiros momentos da indexação assentam na análise de conteúdo que, para Lancaster, “implica a determinação do assunto de um documento e, o mais importante, a decisão de quem o utilizará com maior probabilidade e para que propósito o fará, isto é, antecipar-se às pesquisas para as quais esse documento pode ser relevante [trad. nossa]” (Lancaster, 2002:164). Compreende-se, então, que este processo

Maria José Rodrigues

analítico visa a representação da informação inclusa num documento primário para um documento terciário, este materializado numa linguagem documental, como por exemplo listas de cabeçalhos, tesouros; classificações; índices e resumos, constituindo, assim, a terceira fase da indexação. Nesta perspetiva, não podemos esquecer que um centro de documentação também é um fornecedor de conteúdos, sempre que representa os documentos primários em secundários, como os catálogos e, por sua vez, em documentos terciários, que são as linguagens documentais.

Na verdade, a atividade de analisar o conteúdo de um documento com o objetivo de dele extrair os seus principais assuntos comporta alguns constrangimentos. Para interpretar o conteúdo de um documento, o indexador passa por um processo cognitivo. Sendo a indexação um processo de categorização, torna-se essencial considerar como a atividade do indexador e a estrutura da linguagem de indexação determinam a formação de categorias. Isto porque, durante a indexação, assim como afirma Lancaster, “pode-se cometer vários erros que tendem a ser negativos sobre o rendimento do sistema de recuperação” [trad. nossa] (Lancaster, 2002: 165). Desta forma, segundo o mesmo autor, há cinco tipos de problemas que podem ocorrer. O primeiro são as falhas na análise conceptual, isto acontecerá sempre que o indexador interprete erroneamente o conteúdo do documento. O segundo assenta nas falhas de tradução, ou seja, quando o indexador escolhe termos inadequados para representar o conteúdo informacional. O terceiro erro ocorre por omissão, isto é, sempre que seja omitido um aspeto importante do documento. O quarto lapso que poderá ser cometido é o da falta de especificidade no vocabulário, isto quando não existem termos específicos no vocabulário do sistema e o indexador vê-se obrigado a utilizar outros mais genéricos. O quinto consiste na falta de especificidade da indexação que acontece sempre que o indexador opta por termos mais genéricos do que o assunto concreto do documento, apesar de existir termos mais específicos no vocabulário. Em relação aos conteúdos jornalísticos, tais constrangimentos poderão ocorrer, nomeadamente na indexação de artigos de opinião, podendo falhar a tradução do assunto principal, pois a subjetividade poderá ocultá-lo através de artifícios da linguagem, muitas vezes presente neste tipo de texto.



2.2- Controlo da qualidade

Neste contexto, é pertinente abordar o controlo da qualidade do ato de indexar que se relaciona obrigatoriamente com o indexador, visto se tratar de uma atividade humana, e que abrange a objetividade e a neutralidade. Como determina a NP 3715 (1989), a qualidade e a coerência da indexação dependem de determinados fatores, tais como a competência do indexador e a qualidade dos instrumentos de indexação.

Compreendendo que a qualidade da indexação se verifica tanto no momento de armazenamento da informação como no da pesquisa, a mesma depende de vários fatores. Um deles será a imparcialidade do indexador, para se obter coerência, pois a competência do profissional da informação na análise do conteúdo informacional do documento é crucial, esperando-se que seja conhecedor da área que indexa, embora não necessariamente especializado pois poderá cometer excessos na sua análise já que a tendência, neste caso, será a de interpretar em demasia.

Claro que o utilizador não poderá ser esquecido neste processo de qualidade, visto que a indexação visa a recuperação de uma pesquisa, a qual depende do grau de exaustividade e de especificidade, antes explanadas, e, conseqüentemente, da relação entre o grau de precisão e de revocação.

Precisão é a capacidade do sistema em excluir a recuperação de documentos inúteis, concretiza-se em separar os registos relevantes dos não relevantes, apontando-lhe uma eficácia na pesquisa. Ribeiro (1996: 84-86) determina-a com o seguinte quociente:

Sendo:

a= #A= nº de documentos relevantes recuperados

b= #B= nº de documentos não relevantes recuperados

c= #C= nº de documentos relevantes não recuperados

d= #D= nº de documentos não relevantes não recuperados

A taxa de precisão é:

$$\frac{a}{a+b} = \frac{\text{nº de documentos relevantes recuperados}}{\text{nº total de documentos recuperados}}$$

Já a revocação traduz-se na capacidade de recuperar documentos úteis. Assim como afirma Lancaster (2002), na maioria das situações, o utilizador quer e espera que o



sistema recupere documentos relevantes, conferindo uma eficiência na pesquisa. Ainda Ribeiro (1996: 84-86) apresenta a fórmula para a exprimir quantitativamente.

A taxa de revocação é:

$$\frac{a}{a+c} = \frac{\text{n}^\circ \text{ de documentos relevantes recuperados}}{\text{n}^\circ \text{ de documentos relevante existentes}}$$

Ora, estes parâmetros de recuperação de informação estão intimamente ligados à forma como se indexa, nomeadamente à taxa de especificidade e exaustividade dos termos que representam o assunto do documento. Daí nunca se dissociar o ato do indexador com o do utilizador, havendo mesmo necessidade de existir contacto entre ambos. Aliás, assim como afirma Lancaster, “a exaustividade e a precisão tendem a ser inversas; tudo o que aumenta a exaustividade reduz a precisão; o que aumenta a precisão diminuirá aquela” [trad. nossa] (Lancaster, 2002: 153).

Ribeiro (1996: 84-86) acrescenta outras medidas de qualidade que podem ser aplicadas na avaliação de um sistema de indexação. Uma é complementar da revocação, que é a “taxa de silêncio, sendo a proporção de documentos relevantes não recuperados”, calculando-se da seguinte forma:

A taxa de silêncio é:

$$\frac{c}{a+c} = \frac{\text{n}^\circ \text{ de documentos relevantes não recuperados}}{\text{n}^\circ \text{ total de documentos relevantes existentes}}$$

A mesma autora adverte ainda que “a determinação da taxa de revocação e, logo, da taxa de silêncio, carece de um requisito prévio, que é o conhecimento do número de documentos relevantes existentes, relativos a cada questão colocada ao sistema” Ribeiro (1996: 85).

Outra medida é a “taxa de ruído”, que é a proporção de documentos não relevantes recuperados. Esta medida é complementar à taxa de precisão, calculando-se com a seguinte fórmula:



A taxa de ruído é:

$$\frac{b}{a+b} = \frac{\text{n}^\circ \text{ de documentos não relevantes recuperados}}{\text{n}^\circ \text{ total de documentos recuperados}}$$

Outra taxa proposta pela mesma autora é a “taxa de irrelevância” que exprime a proporção de documentos não relevantes face à quantidade de documentos não relevantes existentes, calculando-se do seguinte modo:

A taxa de irrelevância é:

$$\frac{b}{b+d} = \frac{\text{n}^\circ \text{ de documentos não relevantes recuperados}}{\text{n}^\circ \text{ total de documentos não relevantes existentes}}$$

Na verdade, esta medida é tão importante quanto a revocação, pois a eficácia de um sistema de recuperação de informação é considerada quanto maior for a taxa de revocação e menor for a taxa de irrelevância.

Outra medida apresentada é a “taxa de generalidade”, definida como o número médio de documentos relevantes existentes face a cada questão, calculando-se da seguinte forma:

A taxa de generalidade é:

$$\frac{a+c}{a+b+c+d} = \frac{\text{n}^\circ \text{ total de documentos relevantes}}{\text{n}^\circ \text{ total de documentos existentes}}$$

A autora alerta que a “precisão deverá ser considerada tendo em conta os valores da taxa de generalidade, pois alterações na quantidade de documentos existentes influem nos resultados desta taxa e tornam mais expressivos os valores da taxa de precisão” (Ribeiro, 1996: 86).

Na verdade, o indexador é um leitor profissional, ou seja, assim como afirma Lucas, “o seu instrumento de trabalho é a leitura. É lendo que ele codifica, classifica, indexa, atribui palavras-chave” (Lucas, 1996: 6). Para a mesma autora, o documentalista possui “a competência de saber ler, decifrar criticamente os textos, ler com atenção, de uma maneira bem informada” (Lucas, 1996: 31). Reforça que o “tempo de leitura é somente o suficiente para saber de que trata o texto, não exigindo nada de reflexão, ou de buscar



compreendê-lo, nada de acumular conhecimento. A leitura deve avançar sempre com o fim de extrair do texto o que é útil, o que é produtivo para o utilizador” (Lucas, 1996: 75). Ainda reitera que “o indexador como sujeito-leitor não deve interpretar, a sua leitura deve ser literal, apreendendo o conteúdo do texto e produzindo representações do mesmo, dando-lhe unidade através de palavras-chave” (Lucas, 1996: 95). Ora, nesta perspetiva, o indexador também pode, de uma forma indireta, comentar ao eleger descritores que apresentem julgamentos, trabalho esse que não se espera do profissional da informação. A este propósito, Lucas afirma que “esta é a posição incómoda dos funcionários da memória: de um lado o risco de impor a sua leitura como a leitura de todos, de outro, o de transformar, em pura abstração, sentidos de que se apagaria a memória” (Lucas, 1996: 78). De facto, este carácter interpretativo do indexador poderá pôr em causa a qualidade da indexação e a coerência que lhe deverá ser inerente. Tal poderá acontecer na indexação de conteúdos jornalísticos. Conhecendo o seu potencial informativo, é fácil inculcar uma opinião na escolha de descritores em textos cujos temas são muito sensíveis, como por exemplo, a política ou criminalidade.

2.3- Indexação manual, automática e semiautomática

Esta abordagem não ficaria completa se não se incidisse na forma como o processo de indexação é feito, ou seja, se pelo indexador ou por sistemas automatizados. Quando a indexação é feita pela ação do homem denomina-se indexação manual ou intelectual. Quando é realizada por programas de computador, sempre com o texto em formato digital, ocorre a indexação automática. Lima e Boccato (2009: 6) estabelecem a distinção entre três tipos de operações, afirmando que a indexação realizada por um indexador humano é a denominada por indexação manual; por um programa de computador constitui a indexação automática e, por fim, por um programa de computador e posteriormente revista por um indexador humano será a indexação semiautomática.

A NP 3715 (1989), no seu ponto onde aborda o objetivo e campo de aplicação da mesma, sublinha que as técnicas propostas podem ser empregues em qualquer organismo onde a análise dos documentos e a expressão do seu conteúdo sejam feitas por indexadores, não se aplicando aos organismos que utilizam técnicas de indexação

Maria José Rodrigues

automática “nas quais os termos de um texto são organizados em conjuntos ou classes, segundo critérios que podem ser aplicados por computador como, por exemplo, a frequência da sua utilização e/ou adjacência no texto, ainda que a finalidade deste sistema seja a mesma” (NP 3715, 1989: 3). Percebe-se aqui a necessidade de distinguir o trabalho feito pelo documentalista no processo da indexação manual do realizado por um computador com a indexação automática. Aliás, eleva claramente o trabalho humano nas suas várias vertentes e fases. Contudo, há cada vez mais apologistas para que a indexação automática entre nos procedimentos de vários centros de documentação, como auxiliar do indexador, constituindo, desta forma, uma indexação semiautomática.

Como já foi referido anteriormente, a indexação implica um processo intelectual, resultante de um momento analítico em busca dos assuntos contidos num documento, a fim de proceder à sua representação, tendo em conta os interesses do utilizador. Sendo assim, compreende-se a diretriz da NP 3715 ao sublinhar a intervenção humana neste processo. De facto, o computador apenas executa as orientações humanas, mas de uma forma algorítmica. Será que a máquina consegue competir com o intelecto humano quando tiver de determinar qual o assunto ou descritor mais adequado tendo em conta o público-alvo? É claro que um documento poderá ser representado de diferentes formas tendo em conta os objetivos do organismo, os seus utilizadores e os seus indexadores. Na verdade, não existe uma única forma correta de indexar, tem é de ser feita com coerência e consistência. Neste aspeto, o computador poderá concretizá-lo, mas nunca terá a capacidade de colocar questões e refletir, pois nada substitui esta dimensão humana. Contudo, a ação automatizada poderá servir de auxiliar numa primeira fase da indexação, ao efetuar um primeiro levantamento de assuntos contidos no documento.

A indexação automática estabelece uma interação com a bibliometria, pela aplicação das suas leis e princípios na seleção de descritores, nomeadamente através das primeira e segunda leis de Zipf e do ponto de Transição (T) de Goffman. Em 1948, George Zipf formulou duas leis sobre a distribuição de palavras num texto: *Primeira Lei de Zipf* e a *Segunda Lei de Zipf*. A primeira opera sobre as palavras com alta frequência num texto e a segunda sobre as que têm baixa frequência. Esta, por sua vez, foi aperfeiçoada por Booth, conhecida pela *Lei de Zipf-Booth*. Entretanto, Goffman sugeriu a existência de um ponto (T) onde haveria a transição das palavras de alta frequência para as palavras

Maria José Rodrigues

de baixa frequência. Neste ponto, estariam as palavras representativas do conteúdo do documento. A partir daqui, constatou-se a possibilidade de se aplicar uma lei bibliométrica como instrumento de indexação em sistemas de informação. Mamfrim (1991) considerou mesmo que o vocabulário existente no texto deveria ser a base para a análise do seu conteúdo, sendo esta a melhor forma de recuperá-lo, mudando, assim, a linha de atuação no ato da indexação, já que é dada uma maior importância à linguagem natural em detrimento da controlada. Desta forma, pode-se afirmar que a lei bibliométrica de Zipf é essencial para o processo da indexação automática.

Mamfrim (1991), num dos seus estudos, observou que a distribuição de palavras num texto vai ao encontro das distribuições bibliométricas em geral, concentrando-se na máxima 'poucos com muito e muitos com pouco'. Isto quer dizer que poucas palavras ocorrem muitas vezes enquanto muitas palavras ocorrem poucas vezes ao longo do texto, tendo ficado confirmado que a porção intermédia da distribuição das palavras constitui a área de excelência para a localização das palavras de conteúdo semântico de um texto. Até porque essas mesmas foram encontradas no tesouro da área, sendo então reconhecidas como parte integrante da sua terminologia. Concluiu que a indexação derivativa utilizada reduz incertezas, na medida em que retira do próprio texto palavras para a indexação, ou seja, indexa com as palavras do autor, diminuindo o hiato entre a linguagem natural e a controlada. A autora reforça a ideia que o vocabulário existente no texto deve ser a base para a análise do seu conteúdo, sendo esta a melhor forma de recuperá-lo. Daí a lei bibliométrica de Zipf ser essencial para este processo.

Ainda segundo a mesma autora, a indexação automática consiste na mecanização do processo de indexação, no todo ou em parte, visando estabelecer rotinas que reduzam a interferência da subjetividade do indexador, tanto na análise do documento, quanto na seleção dos termos significativos. Ora, a indexação automática acontece quando sistemas informáticos são utilizados para substituir a indexação manual realizada pelo indexador, ou seja, sem a intervenção direta do homem. Tal contribuirá, de alguma forma, para o controlo de qualidade que a própria Norma 3715 aponta como essencial, que consiste, entre outros aspetos, na exigência da total imparcialidade do indexador, no intuito de se obter a coerência necessária. O facto da referida norma não se aplicar aos organismos que utilizam técnicas de indexação automática, reforça a ideia da necessidade de objetividade na identificação de descritores. Ora, tratando-se de uma



seleção automatizada de conceitos e termos, através da frequência de palavras, concluímos que o perigo da parcialidade desaparece. Porém, outras fragilidades surgem.

Lapa e Corrêa (2011) propõe e avalia um método para seleção de termos relevantes como descritores na indexação automática de teses e dissertações, com a aplicação das primeira e segunda leis de Zipf e o Ponto de Transição (T) de Goffman, cujos focos de estudo são as palavras em textos curtos, isto é, 10 resumos de teses e dissertações. Para tal, utilizou uma ferramenta auxiliar, as listas de *stopwords* ou *stoplist*, que têm como objetivo eliminar grupos de palavras irrelevantes para o assunto do documento, isto é, excluir as palavras sem conteúdo semântico. No seu estudo, faz a comparação das palavras-chave atribuídas pelo autor e as que foram recuperadas na indexação automática. Conclui que, na generalidade, a aplicação da indexação automática gera bons resultados, atribuindo a baixa frequência das palavras-chave no resumo à má escolha dos autores para representar o conteúdo das teses e dissertações. Ficou concluído que bons descritores aumentam o coeficiente de precisão, facilitando a recuperação da informação. Quanto melhor for a indexação maior é a qualidade da recolha de informação. No entanto, alguns conceitos representados por letras isoladas, como unidades de medida e nomes científicos, não foram considerados, tendo sido colocados na lista de *stopwords*, o que conferiu uma fragilidade no sistema de indexação.

Café e Bräscher (2008) alertam para a necessidade de padronizar tanto a descrição física dos documentos como a de conteúdo, para que a recuperação de informação seja eficiente. Isto porque consideram que a padronização aumenta a precisão da recuperação de itens relevantes, já que relembram que o objetivo essencial do processo de documentação é a individualização de um item. Um documento vale pela seleção e individualização desses itens que os transforma em valores de excelência no circuito da informação. Desta forma, uma boa descrição de conteúdo é fundamental para o acesso a itens específicos que se encontram armazenados. O mesmo problema é apontado na análise de conteúdo, também segundo Café e Bräscher (2008), nomeadamente na técnica lógico-semântica, temática ou categorial que conduz ao processo de indexação. O objetivo será inferir as intenções de quem produziu a mensagem de um documento e compreender o seu conteúdo. Para tal, faz-se uma análise quantitativa da frequência dos termos utilizados para depois realizar-se uma análise qualitativa, ou seja, a interpretação



do texto. Daqui extraem-se termos da linguagem natural que representam os assuntos visados. No entanto, se estes termos não forem controlados e transformados em descritores, poderá ocorrer a sinonímia tão indesejada na recuperação da informação. Ora, a aplicação da Lei de Zipf não ultrapassa estas ambiguidades, visto atuar sobre a linguagem natural, não reconhecendo casos de sinonímia, homografia e termos compostos. No intuito de serem ultrapassados esses constrangimentos, só a utilização de instrumentos de indexação de tipo combinatório (tesauros) e do tipo categorial (classificações) é que poderão normalizar as palavras oriundas da linguagem natural do texto para uma linguagem controlada. E só a indexação manual ou intelectual poderá fazê-lo, como se de um filtro se tratasse. Conclui-se, então, que não deixa de existir uma intervenção humana durante o referido processo de uniformização. Para além disso, depararam-se com a dificuldade da aplicação da lei de Zipf, principalmente no problema dos *quase sinónimos*. Estes conceitos, mesmo que relacionados, têm significados diferentes, mediante o contexto no qual estão inseridos, como é o caso de algumas palavras traduzidas. Também encontraram o problema da flexão das palavras, já que o mesmo nome no singular e no plural é contabilizado como distintos, acontecendo o mesmo a um verbo conjugado em diferentes tempos. Para esta redundância, as autoras apresentam a solução de reduzir a palavra ao seu radical. Afirmam, também, que é necessário que a palavra passe a ser delimitada como vocábulo e encarada como uma unidade de léxico. Ora, as leis de Zipf são referentes à extensão do texto, isto é, a verificação é feita pela enumeração das palavras e não de vocábulos. Daí serem contabilizados todas as flexões, o que provoca ruído na análise de conteúdo. Assim conclui-se que, apesar da aplicação das leis de Zipf terem a vantagem da rapidez e diminuição de recursos, apresentam algumas desvantagens, como a dificuldade em medir a representatividade da palavra e a falta de controlo da ambiguidade semântica, ficando a recuperação comprometida pela fraca precisão e revocação.

Desta forma, percebe-se que a indexação manual ou intelectual ajudará a automática, atuando como um filtro na clarificação das suas fragilidades. Em contrapartida, a indexação automática ajuda a manual logo na sua fase inicial, ou seja, na eleição dos assuntos através da frequência das palavras contidas no texto. Neste sentido, é notório que existe uma relação de reciprocidade entre a indexação manual e a indexação automática, o que explica o interesse da aplicação das duas num organismo, denominando-se como indexação semiautomática. De facto, o problema da

Maria José Rodrigues

subjetividade resultante de qualquer atividade humana, como é o caso da indexação intelectual ou manual, poderá suscitar algumas incoerências, daí a indexação automática poder colmatar algumas indecisões do indexador ou, pelo menos, ajudá-lo a clarificar as suas dúvidas no que diz respeito à eleição dos assuntos de um documento, caso pretenda fazer uma indexação mais exaustiva. A inconsistência da indexação manual poderá ser um entrave na recuperação da informação. Por exemplo, um mesmo documento ser indexado com assuntos diferentes em organismos diferentes, assim como o mesmo organismo indexar um documento distintamente em momentos diferentes. Este problema será atenuado caso o texto, em formato digital, seja submetido a uma análise quantitativa para depois passar para a reflexão humana. Na verdade, os princípios da indexação automática são os mesmos dos do ser humano no momento da indexação manual, ou seja, começa por fazer uma avaliação da frequência das palavras dentro do texto assim como a sua posição no mesmo, como serve de exemplo o título, as legendas ou um resumo e, por fim, o contexto desse documento. Este auxílio é essencial nos conteúdos jornalísticos. Atualmente, esses documentos surgem nos arquivos e centros de documentação já em formato digital, sejam os que são publicados em suporte papel ou *online*. Sendo assim, e dada à grande quantidade documental, a indexação automática ocupa um lugar crucial na primeira fase de indexação ao recuperar os termos mais relevantes, normalmente localizados nos títulos, legendas ou *lead*. Assim, o documentalista nestes serviços poderá, desde logo, iniciar uma análise de conteúdo com uma base informacional, que terá de ser revista para a execução das fases seguintes da indexação.

A indexação semiautomática é a mais equilibrada visto que o indexador pode escolher alguns termos atribuídos automaticamente pela frequência de palavras no texto, ajudando-o na sua seleção, porém nunca dispensando este trabalho intelectual. A indexação automática isolada sem qualquer supervisão humana retira a essência do que realmente é a indexação, ou seja, a análise de conteúdo nas suas diferentes vertentes do intelecto humano. A máquina não identifica metáforas, analogismos ou ironias, figuras estas que servem de exemplo para a importância da intervenção humana neste processo. Sabemos que, em distintos conteúdos jornalísticos, o uso de metáforas, entre outros estilos, são apresentados, mesmo em títulos, sendo muitos destes constituídos por citações, principalmente em entrevistas. Ora, as palavras do entrevistado, escolhidas para títulos ou legendas não estarão necessariamente relacionadas com o assunto



principal da entrevista, daí a necessidade da intervenção humana. Na verdade, a indexação automática não alcança o nível de desempenho obtido pela indexação intelectual, porém esse processo reduzirá a carga de trabalho dos indexadores ao realizar uma atribuição preliminar.

3. Política de Indexação

A indexação é uma operação delicada, já que depende de variáveis subjetivas definidas por um indexador humano. Os assuntos são representados por palavras, que muitas vezes carregam diferentes significados em contextos diversos. Por esse motivo, é necessário que o processo de indexação seja desenvolvido a partir de um método que não interfira na intenção do autor, preservando o contexto do documento.

Desta forma, para orientar a indexação dos documentos é necessário estabelecer-se uma política de indexação que reflita a filosofia da organização onde está inserida. Para Araújo, “a política de indexação é um conjunto de diretrizes registadas, que tem a finalidade de nortear o processo de representação da informação, uma das partes mais importantes do tratamento da informação” (Araújo, 2010: 18). Ora, para estabelecer esta política de indexação é necessário adequá-la às características e objetivos do centro de documentação, assim como às necessidades dos utilizadores. Na verdade, o objetivo principal da definição de uma política de indexação é a satisfatória recuperação de informação, ou seja, a capacidade de revocação e de precisão do sistema e os níveis de especificidade e exaustividade, adequados às necessidades dos utilizadores. Para Strehl (1998: 5), os elementos que compõem uma política de indexação são a “cobertura de assuntos” tendo em conta as áreas a serem tratadas; o “processo de indexação” com a definição de variáveis relativas aos níveis de exaustividade e especificidade e a sua capacidade de revocação e precisão; a “estratégia de busca” com a decisão do acesso à base de dados seja pelo documentalista ou pelo utilizador; o “tempo de resposta do sistema” pela identificação do tempo na recuperação da informação; a “forma de saída” tendo em conta a forma de apresentação das informações recuperadas e, por fim, a “avaliação do sistema” com o objetivo de aferir o nível de satisfação dos utilizadores.

No caso dos arquivos e centros de documentação dos meios de comunicação social, esta política de indexação estará intimamente ligada à política editorial da publicação. Na verdade, o que é delineado para ser editado vai ao encontro dos objetivos jornalísticos,

Maria José Rodrigues

principalmente para o público-alvo a que se destina. Assim, a indexação dos conteúdos jornalísticos terá de se justapor às necessidades editoriais e dos jornalistas, que serão os principais utilizadores destes arquivos. Assim, a indexação dos conteúdos de uma publicação generalista será certamente mais abrangente do que a de uma especializada, e essas variáveis deverão estar registadas para haver coerência e consistência. Ainda neste contexto, os meios de comunicação social interferem na construção da opinião pública com a suas opções editoriais, fazendo com que os seus arquivos espelhem essa situação, tanto no momento da indexação como na construção da linguagem documental a utilizar. Na verdade, a mudança de paradigma comunicacional, que assenta na evolução da massificação para a individualização, também poderá estar incutida nas opções da política de indexação, isto porque o facto da exigência dos leitores ser maior fará com que seja necessário uma maior especificidade na escolha dos termos de indexação, para que assim haja uma maior eficiência na recuperação de informação com o aumento da taxa de precisão.

A política de indexação deve ser constituída de estratégias que permitam o alcance dos objetivos de recuperação do sistema de informação. Essas mesmas estratégias deverão ser registadas por escrito para que haja clareza das decisões a tomar, independentemente do indexador. Só assim poderá haver continuidade e aperfeiçoamento ao longo dos anos e de uma forma coerente. Esse registo deverá constituir um manual de indexação, que é um instrumento de trabalho orientador dos princípios de indexação instituídos pelo sistema de informação para que haja consistência na indexação. Este documento assume bastante importância no processo de indexação visto que poderá combater desajustes e subjetividades, devendo ele uniformizar procedimentos ao indicar a ordem lógica de etapas a serem seguidas para a análise de assuntos e ao fornecer as regras e diretrizes para o trabalho do indexador. Desta forma, o manual será um dos meios pelo qual a política de indexação de um sistema de informação poderá ser observada.

Segundo Fujita e Rubi (2006: 13), pode-se considerar que um manual de indexação deve ser uma condensação de três tipos:

- a) Manual de operação ou procedimentos: pois ele deve descrever a atividade de indexação, dar instruções sobre a realização dessa tarefa e proporcionar métodos que possibilitem sua execução de maneira uniforme;



Maria José Rodrigues

- b) Manual de política: o manual deve descrever, de maneira geral e filosófica, as políticas a serem seguidas pelos indexadores no momento da indexação (...);

- c) Manual de organização: deve servir como um repositório das experiências acumuladas dos indexadores mais antigos, a serem aproveitadas para facilitar o treinamento dos mais novos, podendo, com isso, constituir-se num manual de consultas.

Contudo, percebe-se que os dois primeiros serão aqueles que mais se poderá encontrar, pelo facto de terem um campo de ação mais objetivo e prático, ao contrário do terceiro que implica impressões subjetivas do indexador, embora não menos interessantes e úteis, porém mais trabalhosas, acabando por ficar no domínio do que seria ideal.

4. Linguagem natural vs documental

Não é possível dissociar o processo de indexação com as linguagens que representam os conceitos inclusos num documento e destinadas a recuperar essa mesma informação. Ora, após os primeiros momentos de indexação, que consistem na análise do documento, definição do seu conteúdo e seleção dos conceitos representativos, urge a representação desses mesmos conceitos por termos de indexação pertencentes a uma linguagem, seja ela natural ou documental.

A linguagem natural é aquela que está presente nos documentos primários que são submetidos à análise documental, isto é, a que se usa sem qualquer premeditação no quotidiano comunicacional. Na verdade, é sinónimo de discurso comum, constituindo a base para a análise de conteúdo. A linguagem documental é construída a partir da natural pois recolhe dela termos para representar, sob forma condensada, o conteúdo documental. Desta forma, a linguagem documental intervém na terceira fase do processo de indexação, pois trata-se da representação em termos ou descritores dos assuntos eleitos nas duas primeiras fases.

Para aprofundar o conceito destas duas linguagens, importante será também perceber as suas diferenças. Então, se a linguagem natural é usada na comunicação imediata, a linguagem documental é aplicada com um intuito muito específico de comunicação, já que é um código unívoco e estereotipado, controlado e normalizado. Na linguagem natural coexistem diferentes significados para um só significante ou diversas palavras sinónimas. Em contrapartida, a linguagem documental exerce um controlo léxico que impede a utilização de significantes distintos para o mesmo significado. Sendo assim, a linguagem documental tem como principal função organizar e construir o acesso ao documento de forma a facilitar a recuperação da informação.

Claro está que se um organismo estabelecer na sua política de indexação uma indexação meramente automática, estará a servir-se da linguagem natural, pois irá extrair do documento primário os conceitos a partir das palavras mais relevantes, tanto pela frequência como pela sua posição no texto. Porém, se optar por uma indexação manual ou semiautomática, recorrerá a uma linguagem documental que representará a natural contida no texto, de forma a contornar problemas de sinonímia ou homonímia.

Maria José Rodrigues

No caso de arquivos ou centro de documentação dependentes de um meio de comunicação social da imprensa escrita, o recurso à linguagem natural para indexar os conteúdos jornalísticos é perfeitamente possível tendo em conta a posição da palavra no texto, visto estes conteúdos serem estruturados com títulos e legendas. No entanto, como já foi anteriormente referido, existe a necessidade de recorrer a uma linguagem documental para combater situações de ambiguidade que a linguagem natural jornalística também apresenta. Mas o mais importante na construção e utilização de uma linguagem documental num arquivo que indexa conteúdos jornalísticos será o facto dessa mesma linguagem constituir também um conteúdo, já que esses serviços também são fornecedores de conteúdos. Na verdade, se analisarmos um tesauro construído a partir de conteúdos jornalísticos, compreendemos as áreas de conhecimento que abrange e o nível de especificidade que poderá atingir, percebendo-se o perfil informacional desse centro de documentação. Para além disso, a linguagem documental deste tipo de arquivos terá de refletir o potencial informativo que os conteúdos jornalísticos apresentam, como atrás foi descrito. Isto porque os termos eleitos como descritores são a representação da memória que acarretam, pois só assim a mesma poderá ser recuperável. No momento da indexação, o indexador deverá ter sempre esse valor presente numa perspetiva de futuro. Para além disso, a indexação deste tipo de documentos será apoiada numa linguagem jornalística que apresenta características muito particulares, tendo a linguagem documental de a representar também. Sendo assim, o indexador deverá estar familiarizado com a estrutura linguística própria do jornalismo para que a escolha dos termos seja a mais adequada às necessidades informativas dos principais utilizadores que são os jornalistas. Indexar um texto que segue o modelo jornalístico anglo-americano será mais fácil do que um que reflete as características do jornalismo à francesa, isto porque o primeiro apresenta uma linguagem mais objetiva, cujos factos são a prioridade informativa, o que fará com que a eleição do assunto a representar seja mais clara. Se seguir o segundo modelo, a análise de conteúdo terá de ser feita com mais cuidado, sendo a utilização da linguagem documental essencial para eleger os descritores que melhor representem assuntos que estão camuflados por uma escrita mais literária e com alguns artifícios estilísticos. A forma como esses conteúdos são divulgados também poderá ser alvo de atenção para o indexador, já que deverá aproximar o texto à realidade, independentemente dos interesses de vária ordem inerentes à publicação. Esses interesses, que muitas vezes influenciam a democratização da divulgação, não deveria transparecer na eleição dos



assuntos a representar, o que nem sempre acontece, pois a tendência é dar um maior relevo a um conteúdo que teve um grande impacto informacional e uma menor atenção a outro que passou mais despercebido, independentemente do seu valor de memória. Contudo, entende-se que a indexação destes conteúdos tem como principal objetivo a base informacional para a elaboração de outros conteúdos, daí se esquecer, muitas vezes, da sua importância histórica.

4.1- Linguagem documental: livre ou controlada

A linguagem documental livre ou não controlada é aquela em que toda a palavra é recuperada, incorpora novos termos dos documentos primários sem, contudo, relacioná-los, como é o caso das listas alfabéticas. Daqui pode-se associar os sistemas de busca da web, cuja técnica de pesquisa abrange a totalidade do texto sem que haja qualquer controlo e relação dos termos. A busca é realizada a partir do texto completo o que lhe confere lentidão, tendo como vantagem o baixo custo, pois não necessita de especialistas de indexação. Porém, apresenta a desvantagem do facto da recuperação de informação produzir altas taxas de ruído e silêncio, o que provocará perda de informação. Também servem de exemplo as listas de palavras-chave e a de descritores livres. As primeiras são palavras significativas extraídas de forma automática pelo computador ou pelo indexador, do título, do resumo ou de outros elementos paratextuais. Apresentam os inconvenientes de favorecer a ambiguidade semântica como a sinonímia ou polissemia e possibilitam a existência de milhares de palavras-chave, o que não é confortável. As listas de descritores livres são a representação de conceitos significativos por palavras ou expressões, de forma intelectual, sem verificação da sua existência, num instrumento estabelecido *a priori*. Estas já possuem regras de critérios de apresentação dos termos, não contemplando formas verbais isoladas ou adjetivos, diminuindo, desta forma, o número de descritores.

A linguagem documental controlada é uma linguagem codificada, fechada e nominativa, pois define todos os termos que podem ser aplicados para representar o conteúdo dos documentos por áreas de conhecimento, fazendo com que a recuperação de informação seja mais rápida e eficaz. Desta forma, tem como vantagem a redução dos níveis de ruído e silêncio, permitindo ampliar as buscas de informação, já que possibilita a relação dos termos, dos mais gerais com os específicos. Contudo, apresenta a inconveniência de ser mais cara pois requer especialistas na sua elaboração e



atualização, assim como o possível desconhecimento para o utilizador em usá-la, pois ter-se-á de utilizar os mesmos termos que foram escolhidos na indexação. No entanto, esta desvantagem logo se transforma em qualidade, na medida em que proporciona ao utilizador um ponto de pesquisa em vez de dois ou três, tornando esse momento mais prático e objetivo. Para tal, servem como exemplo as classificações, as listas de cabeçalhos e os tesouros. Estas linguagens são construídas *a priori*, isto é, antes de começar a indexar os documentos. Representam de forma unívoca o conteúdo do documento e os termos estabelecem relações entre si, sejam elas de hierarquia e associação, servindo para ultrapassar a ambiguidade da linguagem natural. No âmbito destas linguagens documentais controladas estão também aquelas que se apresentam de uma forma não estruturada, mas cujos termos oferecem medidas de controlo, como é o caso das listas de descritores. Segundo Lancaster, “a indexação tende a ser mais consistente quando o vocabulário utilizado está controlado. Será mais provável que os indexadores estejam de acordo sobre os termos necessários para representar um determinado assunto, se esses são selecionados de uma lista prévia. O mesmo ocorre no processo de pesquisa: será mais fácil identificar os termos apropriados a uma necessidade de informação se são selecionados de uma determinada lista. Portanto, o vocabulário controlado facilita a coincidência entre a linguagem dos indexadores e dos utilizadores” [trad. nossa] (Lancaster, 2002: 22). Ainda o mesmo autor aponta três funções essenciais da linguagem controlada: a primeira será a de reduzir as ambiguidades semânticas através da diferenciação dos homógrafos; a segunda consistirá em melhorar a consistência e a representação do assunto mediante o controlo dos sinónimos e, por fim, a terceira será facilitar a realização das pesquisas amplas ao estabelecer uma estrutura que une os termos relacionados semanticamente.

Desta forma e ainda como o mesmo autor apresenta, poder-se-á reunir os três principais objetivos do vocabulário controlado [trad. nossa] (Lancaster, 2002: 164):

- 1- “Permitir ao indexador a representação consistente dos assuntos dos documentos”.
- 2- “Conseguir a coincidência entre o vocabulário utilizado pelo recuperador e pelo indexador”.



- 3- “Proporcionar meios para que o utilizador possa variar a estratégia de pesquisa para conseguir resultados amplos ou seletivos, mediante as suas necessidades”.

Percebe-se que a linguagem documental controlada compreende custos, como já foi referido. Isto porque a sua construção e manutenção requer esforço, sendo por isso mais cara, já que é mais demorada a seleção de termos de um vocabulário controlado do que a designação livre de palavras-chave. Para além disso, é necessário pessoal qualificado para a sua construção, ao contrário do exigido para a linguagem documental livre. No entanto, no momento da recuperação de informação, esse custo elevado é compensado, pois poupa tempo e esforço. Na verdade, o utilizador obterá resultados mais precisos na sua pesquisa se o vocabulário for devidamente controlado. Porém, em termos de custos e eficácia, opta-se cada vez mais pelo baixo custo no momento de indexação, fazendo com que haja um maior esforço na pesquisa e visualização dos resultados.

4.2- Linguagem documental: pré-coordenada vs pós-coordenada

Para abordar este ponto, é essencial falar da tipologia das linguagens documentais do ponto de vista da coordenação dos termos aplicados na indexação e no momento da pesquisa. Existem dois tipos: classificatório ou categorial e combinatório.

a) Tipo classificatório ou categorial

Estas linguagens documentais são as designadas como pré-coordenadas, pois combinam ou coordenam os termos no momento da indexação, mais propriamente, no da construção da linguagem documental. São exemplo as classificações e as listas de cabeçalhos de assunto. Regem-se sob o princípio da subordinação lógica no interior de um sistema do pensamento, que vai do geral para o particular, do genérico ao específico, subdividindo uma coleção de elementos num número limitado de classes. Estas linguagens surgiram da necessidade em organizar um conjunto de documentos por grandes disciplinas, de forma a ser possível localizar rapidamente a informação desejada, sempre sob o princípio classificatório, para diferenciar os elementos de um conjunto numa estrutura hierárquica (classes e subclasses). Esta estrutura fixa é feita independentemente do conteúdo dos documentos, ou seja, *a priori*. Dado que é uma estrutura hierárquica, há revisões, mas não poderá haver grandes alterações, pois isso implicaria os documentos já indexados. Permitem analisar um



documento quer no seu assunto genérico, quer no específico. Porém, por estar demasiado hierarquizado, pode limitar a recuperação da informação tendo em conta as reais necessidades do utilizador, o que acaba por lhes conferir um carácter restritivo.

As linguagens de tipo categorial ou classificatória são utilizadas, preferencialmente, nas bibliotecas generalistas pois não limitam as suas coleções a um domínio específico. Assim, as linguagens documentais pré-coordenadas tendem a efetuar uma representação do conteúdo mais sintética e generalista. Izquierdo Arroyo e Moreno Fernandez (s.d.) apresentam as consequências da pré-coordenação, explicando que, com estas linguagens, é difícil descrever a multidimensionalidade das relações entre os termos. Para além disso, indicam que como os termos estão dispostos seguindo uma sequência determinada ‘A,B,C,D,E’, tal implica que o primeiro termo (A) seja mais importante do que os outros, fazendo com seja mais difícil combinar os termos no momento da pesquisa.

b) Tipo combinatório

As linguagens de tipo combinatório são as designadas como pós-coordenadas, já que são formadas por termos soltos extraídos da linguagem natural, passíveis de serem combinadas posteriormente, permitindo representar o conteúdo dos documentos de forma analítica e exaustiva. Ora, cada termo guarda o seu valor informacional sem precisar de ser integrado numa formulação pré-determinada, sendo por esse motivo um unitermo. No entanto, podem apresentar listas estruturadas, constituído por vocabulário que vai sendo construído à medida das necessidades da indexação. Partindo do princípio combinatório, um documento com vários conceitos pode ser encontrado através da combinação desses mesmos conceitos, já que a estrutura está preparada de modo a facilitar a procura e a não fixar relações, tendo uma estrutura flexível e mais destinada à pesquisa. São exemplos as listas de descritores; as listas de palavras-chave e os tesauros.

Quando se utiliza uma linguagem combinatória, há uma pós-coordenação no sentido de que os descritores para representar os conceitos presentes nos documentos poderão, no momento da busca documental, ser livremente combinados entre si para formular as consultas que permitirão recuperar esses documentos.



Lancaster (2002) apresenta um exemplo em que um documento é indexado com quatro termos ou classes, comparando a eficácia da recuperação da informação nos dois tipos de linguagens documentais. No caso desse documento estar indexado num sistema pós-coordenado, conserva-se a sua multidimensionalidade da relação entre as quatro classes, ou seja, não é preciso uma ordem de classes, já que todas têm o mesmo peso, permitindo recuperar o documento independentemente da combinação feita na pesquisa. Num sistema pré-coordenado, as classes combinam-se numa determinada sequência durante a construção do índice, o utilizador não poderá combinar livremente as ditas classes, pois não lhe oferece a possibilidade em manipular as mesmas livremente.

Observemos o exemplo dos termos de indexação “IMPRESA” e “CENSURA” organizado em diferentes linguagens documentais e a consequente recuperação:

Linguagem documental	Tipologia	Apresentação	Observações
LISTA DE CABEÇALHO DE ASSUNTO	Categorial Pré-coordenada	Imprensa – Censura – Portugal – 1950-1974 – [Periódicos]	Todos os termos estão ligados numa lógica categorial e coordenados <i>a priori</i> . Aqui, o termo “censura” está subordinado à Imprensa e localizado num espaço e num tempo restritos. A recuperação abrangeria a censura da Imprensa (periódicos) em Portugal entre 1950 e 1974.
CLASSIFICAÇÃO (Classificação Decimal Universal)	Classificatória Pré-coordenada	070 JORNAIS. JORNALISMO. IMPRESA 070.13 Liberdade de Imprensa. Censura. Invasão da privacidade e sua prevenção	Todos os termos estão ligados numa lógica classificatória e coordenados <i>a priori</i> . Aqui, o termo “censura” é uma subclasse de “Imprensa”. A recuperação abrangeria cumulativamente a “Censura”, a “Liberdade de Imprensa” e “Invasão da privacidade e sua prevenção”, mas não a Imprensa. Este termo seria recuperado cumulativamente com “Jornais” e “Jornalismo”.



Linguagem documental	Tipologia	Apresentação	Observações
TESAURO	Combinatória	TG Comunicação TE Meios de comunicação	Todos os termos, apesar de se apresentarem hierarquizados, não dependem dessa estrutura no momento da recuperação, já que cada termo é independente, daí ser pós-coordenado. No entanto, se a pesquisa for direcionada para “Comunicação” (termo mais genérico), todos os documentos indexados por “Imprensa” e “Censura” também serão recuperados, pois lhe estão subordinados.
		TE1 Meios de comunicação social	
	Pós-coordenada	TE2 Imprensa TE Política da comunicação	
		TE1 Comunicação de massas TE2 Censura	

Tabela 1 – Análise comparativa de diferentes linguagens documentais

5. Tesouro: um instrumento de representação da informação

Após uma breve análise comparativa das diferentes tipologias das linguagens documentais, impõe-se a incidência na linguagem pós-coordenada de tipo combinatório, mais propriamente o Tesouro, como instrumento de indexação e recuperação de informação. Isto porque no que concerne aos conteúdos jornalísticos da imprensa, percebe-se que na era digital em que vivemos e, conseqüentemente, na mudança de paradigma comunicacional, o catálogo obriga-se a ser informatizado num sistema integrado e, por isso, um suporte de memória eletrónica. Neste contexto, o Tesouro é o mais adequado para representar as várias áreas de conhecimento que conteúdos desta natureza apresentam, já que é essencialmente um sistema de classificação de assuntos aplicável a um tratamento informático.

Assim como afirma Currás, “não se pode trabalhar num centro de documentação e informação sem dispor de um tesouro adequado e especializado para o centro ou tipo de documentos em questão” [trad. nossa] (Currás, 1998: 7). Daqui depreende-se a importância desta linguagem documental, tanto para o momento de indexação como para a recuperação da informação, assim como indica Cavalcanti, “Tesouro é uma lista estruturada de termos empregada por analistas de informação e indexadores, para descrever um documento com a desejada especificidade, em nível de entrada, e para permitir aos pesquisadores a recuperação da informação que procura” (Cavalcanti, 1978: 27).

A palavra “tesouro” tem origem grega e significa armazenagem, repositório ou tesouro e, por conseguinte, apresentam-se algumas definições numa perspetiva cronológica e histórica. Em 1852, Peter Mark Roget exibiu no seu dicionário intitulado “*Thesaurus of English words and frases*”, publicado em Londres, uma estrutura inovadora, já que aqui as palavras não foram agrupadas segundo a ordem alfabética, mas sim de acordo com as ideias que exprimem e representam, estando organizadas pelo seu significado. Um século depois, o conceito foi recuperado por Helen Brown, mais precisamente em 1957, apontando que a resolução de muitos problemas ao nível da recuperação da informação seria transformar conceitos e as suas relações numa linguagem mais normalizada, isto é, com os sinónimos controlados e as suas estruturas sintáticas simplificadas. Quase duas décadas depois, em 1971, Alan Gilchrist definiu Tesouro como uma lista autorizada de léxicos, diferente de uma Lista de Cabeçalhos de Assuntos, cujas unidades lexicais são



mais pequenas e maleáveis. Nesse mesmo ano, Gernot Wersig apresenta Tesouro como uma lista de termos retirados de textos, coordenados posteriormente para evitar ambiguidades, estabelecendo-se relações hierárquicas, associativas e de equivalência. No ano de 1976, o Tesouro foi definido segundo a sua função e estrutura nos Manuais da Unesco. Assim, quanto à sua função, trata-se de um instrumento de controlo terminológico utilizado para traduzir conceitos de uma linguagem natural para uma controlada. Quanto à sua estrutura, é um vocabulário controlado e dinâmico de termos relacionados e que abarcam um domínio específico do conhecimento. Uma década depois, em 1989, Consuelo Ruiz apresenta o Tesouro como uma linguagem documental combinatória que assegura a união entre o documento e o utilizador, pois a figura do tesouro surge entre o indexador e o utilizador. A Norma Portuguesa 4036 de 1992, correspondente à ISO 2788, define Tesouro como um “vocabulário de uma linguagem de indexação controlada, organizado formalmente de maneira a explicitar as relações estabelecidas *a priori* entre os conceitos” (NP 4036, 1992: 5). Em jeito de súmula, Emília Currás em 1998 define Tesouro como uma “linguagem especializada, normalizada, pós-coordenada, usada com fins documentais, cujos elementos linguísticos – termos simples ou compostos – relacionam-se entre si, sintática e semanticamente” [trad. nossa] (Currás, 1998: 25).

Desta forma, considera-se que um Tesouro deverá ter as seguintes características:

- a) Constituir uma linguagem especializada;
- b) As unidades linguísticas adquirem a categoria de termos, que se relacionam de formas hierárquica, associativa ou equivalência;
- c) Trata-se de uma linguagem documental utilizada em processos de indexação e de recuperação da informação, daí constituir um elo de ligação entre o documento e o utilizador, sendo o indexador o intermediário;
- d) Deve permitir a permanente atualização com a introdução ou supressão de termos;
- e) Tem como objetivo converter a linguagem natural (ambígua e livre) em linguagem controlada.

5.1- Classes de Tesauros

Currás (1998) apresenta uma classificação de Tesauros quanto ao tema e sua apresentação, à língua e à entidade que o constrói, assim como está estruturado na seguinte tabela:

Tema	Apresentação do tema	Língua	Entidade
Gerais	Alfabéticos	Monolíngues	Públicos
Especializados	Sistemáticos (temáticos)	Bilíngues	Privados
Multidisciplinares	Hierárquicos	Plurilíngues	
Monodisciplinares	Facetados		
Principais	Gráficos		
Auxiliares	Estrutura alfabética		
	Estrutura sistemática		
	Estrutura gráfica		
	Macrotesauro		
	Microtesauro		

Tabela 2 – Classes de Tesauros

Todas estas classes apresentadas estão interligadas, sendo algumas dependentes de outras. Por isso, quanto ao **tema**, os tesauros gerais são naturalmente multidisciplinares, sendo muito difícil utilizar somente um monodisciplinar. Aliás, poder-se-á utilizar um tesouro principal juntamente com um auxiliar. No que diz respeito à **apresentação do tema**, os tesauros sistemáticos podem apresentar os assuntos de uma forma hierárquica, assim como distribuídas em facetadas ou, então, dispostos em gráficos. Aliás, a maior parte dos tesauros apresenta-se alfabética, sistemática e graficamente. Um macrotesauro é composto por vários microtesauros relacionados entre si, ou seja, é um tesouro que contém temas diferentes dispostos em várias secções. Um microtesauro abarca um tema especializado, podendo desligar-se do macrotesauro e atuar autonomamente. Quanto à **língua**, um tesouro pode ser construído numa só língua, como também em duas ou mais.



Relativamente à **entidade** que o constrói, esta pode ser pública ou de índole privada, o que poderá alterar a política de indexação e, conseqüentemente, o tesouro correspondente, já que os objetivos desse centro de documentação variará consoante a sua missão. Um centro de documentação público que indexa conteúdos jornalísticos terá certamente um cuidado maior em fazer prevalecer o seu potencial informativo de valor de memória, já que tem a missão de legado. Um arquivo privado desta natureza terá uma preocupação maior no potencial da democratização da divulgação, tendo como objetivo ir ao encontro do que é necessário no imediato, prevalecendo interesses de audiência.

5.2- Classes de termos: composição e relevância

Entende-se que o termo de indexação representa o conceito, que é a matéria-prima da indexação, constituindo uma unidade de pensamento, uma unidade abstrata criada a partir de uma combinação única de características. Segundo a NP 4036 (1992: 5), um termo de indexação é a “representação de um conceito, sob a forma de um termo derivado da linguagem natural, de preferência um substantivo simples ou composto”. Ora, num tesouro, os termos, quanto à sua **composição** ou forma, podem ser simples ou compostos e relacionam-se de forma a permitir a sua combinação num processo de pós-coordenação. Por esse motivo, são flexíveis e a sua atualização é dinâmica e rápida. Para além disso, podem ser acompanhados de qualificadores e notas explicativas ou de aplicação.

O termo simples deverá ser constituído por um nome. O termo composto, ainda segundo a mesma norma, “pode ser decomposto morfológicamente em componentes distintos, podendo cada um ser expresso ou reexpresso por uma palavra susceptível de servir, por si própria, de termo de indexação” (NP 4036, 1992: 5). Esta classe de termos é constituída por um núcleo que designa a classe genérica dos conceitos à qual o termo pertence e por um modificador ou distintivo que tem como objetivo restringir o sentido do núcleo ao especificar uma das suas classes. Por exemplo, o termo composto “Jornalismo desportivo” é constituído pelo núcleo que abrange a grande classe do conceito “Jornalismo”, sendo o adjetivo “desportivo” o modificador, já que especifica o sentido do núcleo. Como são constituídos por várias palavras, sendo uma delas um nome correspondente ao núcleo, os termos compostos são considerados expressões nominais e apresentam-se sob a forma de expressões adjetivas ou expressões prepositivas. Nas primeiras, o nome é acompanhado por um adjetivo correspondente ao

modificador, como por exemplo “Indústria audiovisual” e nas segundas, ao núcleo é acrescentado uma preposição e um nome como distintivo, como serve de exemplo “Profissional da comunicação”. No entanto, como consta na já referida norma NP 4036, “uma pequena proporção de termos compostos não pode ser submetida a este tipo de análise lógica. São termos nos quais um dos elementos possui a função de distintivo, mas não especifica uma subclasse de núcleo” (NP 4036, 1992: 18). É o caso do termo composto “Autoestrada da informação”, pois “informação” não constitui uma subclasse de “autoestrada”, nem esta uma grande classe de informação. Este tipo de termos compostos assumem o nome de “sincategoremático porque não pode servir de indicador de classe de noções à qual se refere o termo inteiro. O termo como um todo deve ser tratado como uma só unidade semântica” (NP 4036, 1992: 18) e, por essa razão, nunca poderá ser decomposto.

Os termos de indexação podem ser agrupados por três categorias tendo em conta as noções que representam:

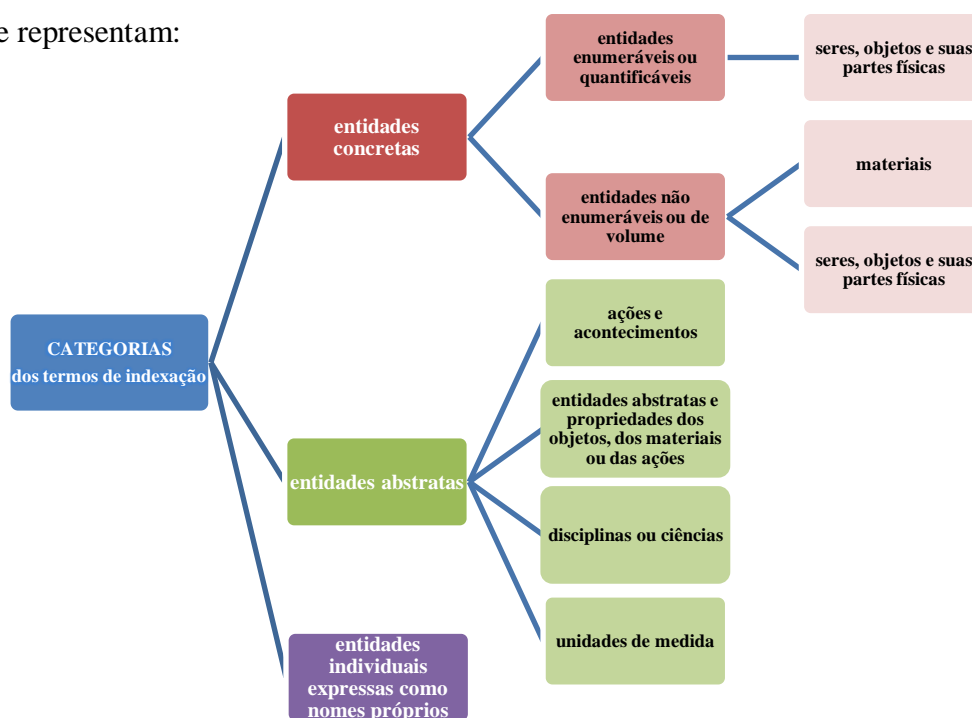


Figura 3 – Categorias dos termos de indexação

Estas categorias interferem na forma dos termos no que diz respeito ao número, ou seja, a utilização do plural e do singular. Assim como determina a NP 4036, “termos que representam entidades enumeráveis (...) devem ser expressos no plural”. Já os termos que representam entidades não enumeráveis ou de volume “como nomes de materiais ou de substâncias (...) devem ser expressos no singular” (NP 4036, 1992: 12). O mesmo

acontecerá às entidades abstratas, tais como fenómenos, propriedades, religiões, atividades e disciplinas. Contudo, poderá ser necessário o uso de um qualificador para casos em que “as formas singular e plural de um termo se reportam a noções diferentes” (NP 4036, 1992: 13). Denote-se que o qualificador é parte integrante do termo, seguindo-o por parêntesis, tendo como objetivo controlar a homonímia; homografia e a polissemia que tanto contribuem para a ambiguidade semântica. Por essa razão, percebe-se que deverá existir um trabalho consistente e coerente no momento da construção do tesauro e do controlo do vocabulário para que a indexação usufrua de uniformidade. Isto porque num tesauro, os termos devem ter um sentido restrito, mesmo que possam ter outros significados, ao contrário do que acontece num dicionário. Por esse motivo, em casos menos explícitos, para além do uso de um qualificador, os termos poderão ser acompanhados por uma nota explicativa ou de aplicação para “limitar e precisar o sentido em que estes devem ser utilizados e assim excluir outros possíveis significados” (NP 4036, 1992: 16), que pode ser a data de adoção do termo assim como a sua fonte no caso de ser um neologismo ou outro tipo de instruções para o indexador. Na verdade, as notas explicativas não fazem parte do termo e, por isso, não são considerados descritores. No entanto, Currás distingue duas funções destas notas, afirmando que “as notas de aplicação que determinam o tema dos homónimos fazem parte do termo” [trad. nossa] (Currás, 1998: 48), ao contrário das que servem apenas para clarificar o seu uso. A partir daqui entende-se que as notas explicativas podem substituir o uso do qualificador, sendo vantajoso sempre que se queira diminuir a extensão do termo de indexação, como pode ser visualizado na grelha que se segue:

Exemplo	Designação
Rádio (aparelho recetor de sons) Rádio (osso) Rádio (estação radiodifusora)	Termo + qualificador = descritor
Rádio NE Aparelho recetor de sons	Termo = descritor + nota explicativa = não-descritor (cfr. NP 4036, 1992:16) Termo + nota explicativa = descritor (cfr. Currás, 1998:48)
Rádio NE Osso	
Rádio NE Estação radiofónica	

Tabela 3 – Uso do qualificador e das notas explicativas

Maria José Rodrigues

Duas categorias de termos de indexação de uma linguagem controlada podem ser definidas quanto à sua **relevância**: os descritores e os não-descritores. O descritor é o “termo que se utiliza na indexação para representar um determinado conceito, por vezes chamado «termo preferencial», já o não-descritor é o termo não-preferencial, “sinónimo ou quasi-sinónimo de um descritor” (NP 4036, 1992: 5), não podendo este ser atribuído a documentos, estabelecendo uma relação associativa com o descritor. A escolha dos termos como descritores deverá ser rigorosa de forma a aplicar coerência na indexação, daí o controlo do vocabulário ser de extrema importância. Por esse motivo, vários aspetos deverão ser levados em consideração durante essa definição. A NP 4036 (1992) propõe alguns critérios na escolha de termos como descritores em detrimento de outros que ficarão como não-descritores. É o caso das diferentes ortografias que um vocábulo poderá conter, devendo-se privilegiar a forma cujo uso corrente é mais generalizado. A propósito deste constrangimento, parece pertinente deixar o alerta para as mudanças lexicais que a língua portuguesa está a sofrer com a implementação do Novo Acordo Ortográfico de 1990. Urge uma adaptação nas bases de dados existentes e, também, nos recursos de autoridade, nomeadamente nos Tesouros. Este processo de renovação e uniformização ortográfica provocará uma mudança na apresentação gráfica de uma mesma palavra, apesar de nada mudar ao nível do conteúdo. Contudo, num sistema de recuperação de informação, é de extrema importância a inclusão das novas grafias através do uso de remissivas, para que não se duplique o mesmo conceito, pois tal irá influenciar a recuperação da informação. Outro critério é o caso de estrangeirismos que ganharam expressão no uso da língua, devendo, por isso, serem eleitos como descritores. Se caso a sua tradução coexistir, deverá ser escolhido como descritor aquele que tiver mais uso. Já se percebe que, por norma geral, deve escolher-se a forma consagrada no uso corrente, como é exemplo os nomes científicos assim como os nomes de lugares e de pessoas físicas e coletividades. Tal regra já não se aplica com os nomes comerciais, o calão ou a gíria.

5.3-Estrutura e Relações básicas

A partir do que foi anteriormente exposto, urge a abordagem das relações básicas que os termos estabelecem entre si. Desta forma, nos Tesouros conhecem-se três tipos de relações entre os termos, assim como é apresentado na NP 4036 (1992):



- a) **Relação de equivalência:** é a relação entre descritores e não-descritores, sempre que exista vários termos para o mesmo conceito, aplicando-se aos sinónimos e quasi-sinónimos. A reciprocidade é expressa por meio das convenções UP (Usado por) e USE.

Exemplos:	Tesouro UP Thesaurus	A forma portuguesa assume-se como descritor por ter mais uso.
	JN USE Jornal de Notícias	O acrónimo assume-se como não descritor visto a forma desenvolvida ser mais reconhecida.
	Acta USE Ata	A forma da nova grafia assume-se como descritor.

- b) **Relação hierárquica:** baseia-se em relações de subordinação, sendo o termo superior representativo de uma classe ou um todo e os termos subordinados partes desse todo. A reciprocidade é expressa pelas convenções TT (Termo de topo); TG (Termo genérico) e TE (Termo específico).

Exemplo:	Meios de comunicação social TE Imprensa TE Radiodifusão TE Televisão	Os termos específicos “Imprensa”; “Radiodifusão” e “Televisão” estão subordinados ao termo genérico “Meios de comunicação social” através de uma relação genérica entre a classe e seus membros.
	Imprensa TG Meios de comunicação social	

- c) **Relação associativa:** inclui relações entre termos que não constituem nem uma relação de equivalência nem hierárquica, contudo mentalmente associados. A reciprocidade é expressa pela convenção TR (Termo relacionado).

Exemplo:	Radiodifusão TG Meios de comunicação social TR Televisão TR Imprensa	Os termos relacionados pertencem à mesma categoria pois contém o mesmo termo genérico.

Sendo assim, os termos num tesauro, quando estão estruturados, são antecidos de abreviaturas e símbolos, que indicam a relação ou a função do termo, assim como é apresentado na tabela que se segue:

abreviaturas	significado	relação/função
NE	Nota explicativa	Indica/contextualiza o sentido do termo, precisando-o. Explicita os sentidos que se excluem.
USE	Use	Antecede o descritor, é usado depois do não-descritor ou termo não preferencial e estabelece uma relação de equivalência.
UP	Usado por	Antecede o não-descritor, é usado depois do descritor ou termo preferencial e estabelece uma relação de equivalência.
TT	Termo de topo	Antecede o nome da classe mais genérica e estabelece uma relação hierárquica.
TG	Termo genérico	Antecede uma noção com um sentido mais amplo e estabelece relações hierárquicas.
TE	Termo específico	Antecede uma noção com um sentido mais restrito e estabelece relações hierárquicas.
TR	Termo relacionado	Antecede um termo associado que não é sinónimo e estabelece uma relação associativa.

Tabela 4 – Abreviaturas dos termos num Tesauro

5.4- Gestão da construção de um tesauro para conteúdos jornalísticos

A utilização de um Tesauro torna-se de extrema necessidade em qualquer centro de documentação ou arquivo, tanto no momento da indexação como, também, para a recuperação da informação. Aliás, com a utilização de sistemas automatizados, esta linguagem pós-coordenada é a ideal, na medida em que cada termo de indexação assume um grau de autonomia que se coaduna com a possibilidade de relacionar, ordenar ou substituir cada termo. E é sobre os termos que é necessário começar a refletir quando se inicia o processo da construção de um Tesauro. O primeiro critério será sempre os objetivos do organismo e seus utilizadores. No caso de um arquivo ou centro de documentação de um meio de comunicação social, os utilizadores são, na esmagadora maioria, os jornalistas. Por esse motivo, os termos de indexação deverão ir ao encontro do jargão jornalístico utilizado, assim como dos objetivos da política editorial. Como já foi referido anteriormente, a política de indexação deverá caminhar

Maria José Rodrigues

lado a lado dos objetivos editoriais, tendo o documentalista de ser conhecedor dos mesmos. Só assim é que o Tesouro a utilizar na indexação dos conteúdos jornalísticos será um verdadeiro instrumento de recuperação da informação. A propósito destas operações documentais, Currás (1998) distingue um Tesouro de indexação de um Tesouro de recuperação de informação, afirmando que “em muitos centros de documentação utilizam-se tesouros oriundos de base de dados e redes de informação externas, dos quais não necessitam de utilizar na sua totalidade. Nesse caso constrói-se e/ou adapta-se à medida das suas necessidades. Este será o tesouro de recuperação de informação, construído com as mesmas regras do de indexação” [trad. nossa] (Currás, 1998: 74). Na verdade, trabalhar com este instrumento elaborado por alguém externo comportará algumas dificuldades, na medida em que o tesouro deverá ser construído tendo em conta o âmbito do(s) tema(s), as necessidades dos utilizadores ou o número de consultas ou pesquisas. Por isso, nesse contexto será sempre obrigatório proceder a adaptações. Então, o ideal é investir na composição de um tesouro de raiz, pois assim irá ao encontro das reais necessidades dos utilizadores e, também, dos indexadores, constituindo, assim, um Tesouro de indexação. Ora, num centro de documentação dependente de um meio de comunicação social o mesmo acontece. Se a indexação e a recuperação da informação dos conteúdos jornalísticos se correlacionam com a política editorial do organismo, então será de todo coerente que o tesouro seja próprio.

Na verdade, um arquivo de um meio de comunicação social trabalha com documentos primários, com uma fonte de proveniência interna, o que confere confiança, e com uma necessidade de recuperação permanente. Sendo assim, dever-se-á construir um tesouro mediante as necessidades concretas de cada caso e, desde logo, definir a que classe pertencerá, ou seja, se geral e multidisciplinar ou, então, especializado e monodisciplinar. Para além disso, pode haver necessidade de fazer mais do que um tesouro, um principal e outros auxiliares. No caso desta tipologia de centros de documentação, tudo dependerá da natureza do organismo. Se for de um jornal generalista, por exemplo, certamente que terá conteúdos abrangentes de várias ordens e, por isso, necessitará de um tesouro multidisciplinar e geral. Contudo, poderá eventualmente recorrer a vários tipos de tesouros, nomeadamente para distinguir termos de assuntos generalistas dos de carácter mais especializado, como os de índole geográfica. Claro que se se tratar de uma publicação especializada em economia, por exemplo, o tesouro será certamente mais específico e monodisciplinar.

Maria José Rodrigues

Relativamente aos termos a eleger, a NP 4036 (1992) apresenta dois métodos de compilação. O primeiro é o método dedutivo, em que se “extraem termos de documentos no decorrer de uma fase preliminar de indexação sem tentar controlar o vocabulário ou determinar as relações entre os termos, até que se tenha recolhido um número suficiente. Todos os termos são, em seguida, examinados” (NP 4036, 1992: 48), sendo o controlo do vocabulário efetuado à medida que se vão estabelecendo as várias categorias. O segundo constitui o método indutivo, a partir do qual “aditem-se no tesouro os novos termos à medida que vão sendo encontrados nos documentos. Efectua-se o controlo do vocabulário desde o início e cada termo, no momento da sua admissão, é colocado numa ou várias categorias anteriormente estabelecidas” (NP 4036, 1992: 48). Com este método, o tesouro é construído partindo do particular para o geral num processo contínuo. Ora, nos centros de documentação que indexam conteúdos jornalísticos, a probabilidade de se construir um tesouro a partir deste último método é maior, visto ser mais óbvio construí-lo continuamente à medida que os vários documentos vão sendo publicados e tratados. Até porque as temáticas abordadas num meio de comunicação social são diversas e abertas, assentes numa mutação constante. O que poderá eventualmente acontecer é a inclusão de termos no momento que o controlo do vocabulário se realiza, procedendo-se, dessa forma, ao método dedutivo.

E é a propósito do controlo de termos que surge a necessidade de os validar. E, para tal, o documentalista tem de considerar dois critérios: o da autoridade literária ou bibliográfica e o da autoridade do utilizador. Assim como afirma Lancaster, “um termo está justificado se aparecer na bibliografia do respetivo tema, ou melhor, se aparecer com a suficiente frequência na literatura, para que seja considerado significativo e útil para a sua recuperação” (Lancaster, 2002:42). Ora, no caso dos conteúdos jornalísticos da imprensa, a palavra toma um relevo significativo, refletindo um dos seus potenciais informativos, o da narratividade, o que faz com que o critério de autoridade literária ou bibliográfica seja essencial na confirmação dos termos a incluir no tesouro. Claro está que o processo de indexação automática pode ajudar a execução do processo por este critério, na medida em que possibilita a leitura dos termos no texto quanto à sua frequência ou, mesmo, a partir dos títulos ou legendas. Porém, será sempre necessária a sua validação intelectual. Por outro lado, o critério de autoridade do utilizador não deve ser esquecido. Isto porque, como atesta Lancaster, “justifica-se a inclusão de um termo num índice caso seja de interesse para o utilizador. Isto é muito importante para

Maria José Rodrigues

estabelecer um nível adequado de especificidade no vocabulário (...) porque os termos que surgem na bibliografia podem ser mais específicos do que aqueles que realmente os utilizadores necessitam” [trad. nossa] (Lancaster, 2002: 43). A atender a este critério para a construção de um tesouro, será necessário proceder a inquéritos, entrevistas ou questionários aos utilizadores, o que implicará um maior tempo para a concretização da linguagem documental, o que nem sempre é desejável. No caso de um arquivo ou centro de documentação de um meio de comunicação social, os utilizadores com maior índice e interesse de consulta são os jornalistas. Por serem próximos, pode ser mais fácil a resposta a este critério, que comporta pertinência na elaboração do tesouro. Na verdade, trata-se de uma interação que já existe informalmente, pois é o jornalista que produz os documentos primários que, por sua vez, chegam ao documentalista para que este os organize em catálogos na linha de documentos secundários. Em contrapartida, para que os jornalistas consigam recuperar a informação contida nas notícias que antes elaboraram, é necessário que o seu colega documentalista também produza bons instrumentos documentais terciários, como é o caso do tesouro. Então, por que não um trabalho colaborativo nesse momento?

Outro aspeto a considerar na gestão da construção de um tesouro é a sua atualização, revisão e manutenção. Isto porque um vocabulário controlado não é estático, pois ele sofre um crescimento ao longo do tempo. Para além disso, alguns termos poderão ficar desatualizados, principalmente aqueles que pertencem às categorias de entidades concretas e entidades individuais expressas como nomes próprios, nomeadamente os que estão relacionados com a ciência exata e materiais, assim como com as novas tecnologias. No caso dos conteúdos jornalísticos o mesmo acontece, já que novos assuntos surgem constantemente na atual era comunicacional. É claro que a necessidade de atualizar os termos de um tesouro aumenta se o grau de especificidade for elevado, daí ser um aspeto a considerar antes mesmo de o começar a construir. No entanto, o aumento do vocabulário deve ser contínuo e basear-se nas autoridades bibliográficas e do utilizador. Lancaster propõe que “se um indexador encontra um assunto que não esteja representado corretamente com o vocabulário existente, pode estabelecer um descritor candidato” [trad. nossa] (Lancaster, 2002: 127). Este descritor candidato será depois analisado para adequá-lo às normas e política de indexação estabelecida no organismo, de forma a inclui-lo no lugar adequado do tesouro. Ainda segundo o mesmo autor “nalguns sistemas, os novos descritores são incorporados durante um período

Maria José Rodrigues

probatório, durante o qual se analisa os seus padrões de utilização (forma e frequência de uso) para rejeitá-lo ou incorporá-lo definitivamente” [trad. nossa] (Lancaster, 2002: 127). O que poderá também acontecer é fazer com que o sistema guarde as palavras da linguagem natural que os utilizadores utilizam nas suas pesquisas para, depois, os indexadores procederem a uma análise e perceberem a pertinência de alguns para a inclusão controlada no tesouro. Na atualização do vocabulário também se inclui a eliminação de termos. Para tal, as estatísticas de utilização poderão ser indicadores úteis, tanto a utilização no momento da indexação como na recuperação da informação. Se um termo não for utilizado numa pesquisa, durante pelo menos dois anos, então será um sinal de que não é necessário ou, então, que existe uma discrepância entre a indexação e os interesses dos utilizadores. Tudo isto tem de ser estudado para que não se cometa erros de forma a provocar elevadas taxas de silêncio ou de ruído. A NP 4036 adverte para estes casos que “se num tesouro se suprime um termo que já tinha sido utilizado para indexação, o mesmo deve manter-se no tesouro com indicação «só para recuperação» e com a data de supressão” (NP 4036, 1992: 49).

Em forma de sùmula, compreende-se que a revisão de um tesouro deve ir ao encontro da utilidade dos termos, tanto no momento da indexação como no da pesquisa. Por isso, ter-se-á de ter em conta se um termo é muito usado ou se caiu em desuso, se apareceram novos conceitos que precisam de novos termos e proceder à sua introdução, ou então, se é necessário alterar as relações entre os termos na medida em que o conteúdo conceitual dos termos também poderá sofrer alterações ao longo do tempo. Este processo de gestão é muito delicado, pois o desuso de termos poderá ser temporal. E isto é muito provável acontecer em conteúdos jornalísticos, visto que se assiste a uma célere mudança de interesses temáticos, havendo, desta forma, a necessidade de um grande cuidado nesta atualização.

Interessa sobretudo fixar a importância do tesouro como um documento terciário que reflete a estrutura informacional dos documentos primários, sempre em conformidade com o que os utilizadores necessitam no momento da recuperação da informação. Neste contexto e numa perspetiva metafórica, Lucas afirma que “os tesouros são o passaporte que abre a entrada para o território da memória inscrita nas bibliotecas” (Lucas, 1996: 71). Acrescenta ainda que “a construção dos tesouros apoia-se no corpus discursivo da área que abrange, pretendendo com isso a garantia literária, isto é, a seleção de termos é



Maria José Rodrigues

feita dentre os conceitos mais frequentes” (Lucas, 1996: 72). Sendo assim e ainda segunda a mesma autora “os tesouros cristalizam os conceitos, fixando-os e dando-lhes autoridade”, daí possuírem um “caráter de grade interpretativa, onde cada descritor é a síntese de um amplo espectro de significados” (Lucas, 1996: 72). Por isso, considera-se o tesouro como um instrumento de representação da informação e do conhecimento. E é esta dimensão que nunca poderá ser esquecida pelo documentalista que constrói e gere um tesouro em qualquer biblioteca, arquivo ou centro de documentação.

III – ESTUDO DE CASO: PRÁTICAS DE INDEXAÇÃO DO ARQUIVO E CENTRO DE DOCUMENTAÇÃO DO GRUPO CONTROLINVESTE

Para atingir os objetivos delineados para este trabalho de investigação, foi feito um estudo com base na metodologia de observação direta, *in loco*, no Arquivo e Centro de Documentação do Grupo Controlinveste. Foram realizadas várias visitas às instalações do Porto, entre o dia 13 de abril de 2012 e 23 de julho do mesmo ano, num total de 20 dias, correspondendo a cerca de 60 horas.

Inicialmente, durante o mês de abril de 2012, foi feito um levantamento do funcionamento do sistema integrado e um reconhecimento dos Tesouros existentes.

Nos meses seguintes, em maio e junho, foram feitas pesquisas no sistema do arquivo digital, a partir de descritores existentes nos referidos Tesouros, para então analisar a dinâmica da recuperação de informação e as variáveis das práticas de indexação aplicadas.

Por fim, durante o mês de julho de 2012, foi feita uma análise do uso dos Infraconceitos nas fotografias, já no novo sistema integrado.

Após um processo de maturação e análise dos dados recolhidos, de forma a consolidar e aferir as conclusões deste estudo, foi realizada, no dia 07 de junho de 2013, uma entrevista ao Diretor do Arquivo e Centro de Documentação do Grupo Controlinveste.

Fica então o esquema de trabalho deste estudo de caso:

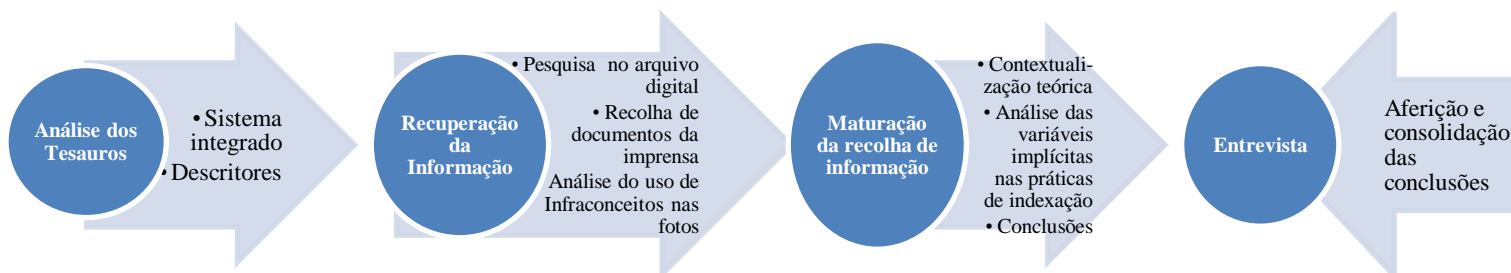


Figura 4 – Estudo de caso: esquema de trabalho

1. Contextualização

1.1- Grupo Controlinveste

O Grupo Controlinveste é um dos maiores grupos de media em Portugal, com presença nos setores da imprensa, rádio, televisão e internet, tendo também participações em empresas de publicidade, comunicação multimédia, produção de conteúdos e design, telecomunicações e desporto. Com origem na Olivedesportos, empresa fundada em 1984, foi crescendo, adquirindo o seu primeiro título de imprensa, o jornal desportivo “O Jogo”. Em 1998, lança a SportTV em parceria com a RTP e a PT Multimédia. Em 2001, cria a Sportinveste Multimedia, em parceria com a Portugal Telecom, empresa essa responsável pela gestão das operações digitais e multimédia dos três principais clubes de futebol portugueses.



Ilustração 1 – Edifício do Jornal de Notícias / Arquivo e Centro de Documentação do Grupo Controlinveste (Porto)

Para além disso, lançou outros canais de televisão, SportTV2, SportTV3 e Sport TV HD. Após a aquisição da Lusomundo Serviços em 2005, o Grupo Controlinveste reuniu uma das maiores ofertas na área dos media, que inclui a SportTV, a TSF e títulos dos maiores jornais portugueses, como o Jornal de Notícias; Diário de Notícias; 24 Horas; O Jogo; Global Notícias; Ocasião; Açoriano Oriental; Jornal do Fundão; Diário de Notícias da Madeira; as revistas Evasões; Volta ao Mundo, para além de uma participação acionista na agência Lusa. Outras áreas de intervenção do Grupo Controlinveste assentam no setor da impressão através das empresas gráficas Funchalense (Lisboa) e NavePrinter (Porto) e da distribuição, com as empresas VASP e Notícias Direct. De uma forma complementar, a Controlinveste têm também várias participações financeiras em sociedades desportivas e empresas de telecomunicações assim como no setor do turismo com a agência de viagens Cosmos. As últimas notícias associadas a este grande grupo prendem-se com a venda da maioria do seu capital a um fundo angolano.

1.2- Arquivo e Centro de Documentação do Grupo Controlinveste: caracterização



O Grupo Controlinveste tem dois centros de documentação, um no Porto e outro em Lisboa, constituindo a Direção de Documentação e Informação (DDI) Global Notícias Publicações, SA (Controlinveste Media – Grupo Controlinveste). O acervo corresponde à produção editorial dos periódicos ‘Jornal de Notícias’; ‘Diário de Notícias’; ‘O Jogo’; ‘24 Horas’; ‘Tal & Qual’; ‘Jornal do Fundão’; ‘Açoriano Oriental’; ‘Ocasião’; ‘Playstation2’ e ‘Volta ao Mundo’.

Ilustração 2 –Edifício do Diário de Notícias / Arquivo e Centro de Documentação do Grupo Controlinveste (Lisboa)

1.2.1 - Arquivo físico

No **Arquivo e Centro de Documentação** do grupo Controlinveste encontram-se arquivados milhões de documentos fotográficos e textuais em papel.

Este espólio é formado por páginas de todos os jornais do grupo, nomeadamente do Diário de Notícias (desde 1864); Jornal de Notícias (desde 1888); 24horas (desde 2004); O Jogo (desde 1998); Jornal do Fundão (desde 2005); Açoriano Oriental (desde 2006); Tal & Qual (de 1984 a 2007) e Global Notícias (desde 2007), para além das revistas semanais e mensais do grupo, como é exemplo Notícias Magazine, Notícias Sábado, Notícias TV, Volta ao Mundo e Evasões. A documentação/informação tratada em Arquivo data de 1927 em Lisboa e de 1957 no Porto.

O acervo da Controlinveste inclui também um conjunto de 300 mil dossiês temáticos e biográficos formados a partir de recortes de jornais e revistas. Esses recortes estão organizados em envelopes que, por sua vez, estão dispostos por personalidades. No Arquivo das instalações do Porto, o acervo é constituído por cerca de 3 milhões de recortes e 3 milhões de fotografias. Em Lisboa, são cerca de 6 milhões de recortes e 8 milhões de fotografias, estas em papel, chapas de vidro e negativos.

Ilustração 3 – Área de recortes do Arquivo e Centro de Documentação do Grupo Controlinveste (Porto)



1.2.2- Circuito documental: jornal / arquivo

O circuito documental jornal/arquivo inicia, naturalmente, na redação através dos sistemas editoriais. Assim que os jornais fecham a sua edição, toda a publicação é enviada para a base de dados do Arquivo com uma indexação baseada em texto livre, por regiões e secções. Depois, é feita uma indexação mais fina pelos documentalistas com base nos tesouros existentes, utilizando não só a parte temática como a onomástica e geográfica. Esta parte textual fica disponível para o mercado no final da manhã, após o tratamento documental. Durante a tarde, os documentalistas dedicam-se ao tratamento de fotografias, assim como à correção dos erros ortográficos e à indexação pela parte temática, de acordo com o texto.

Ilustração 4 – Arquivo e Centro de Documentação do Grupo Controlinveste (Porto)





Maria José Rodrigues

1.2.3- Arquivo digital e sistema de informação

Os documentos são propriedade / oriundos dos meios de comunicação do grupo Controlinveste:

24 HORAS	JORNAL DE NOTÍCIAS	O JOGO
AÇORIANO ORIENTAL	JORNAL DO FUNDÃO	PLAYSTATION2
DIÁRIO DE NOTÍCIAS	NOTÍCIAS MAGAZINE	REVISTAS
RECORTES	NOTÍCIAS SÁBADO	TAL & QUAL
GLOBAL NOTÍCIAS	NOTÍCIAS TV	VOLTA AO MUNDO

A base de dados QUAY, com capacidade para 6 milhões de registos, ultrapassou os seus limites, sentindo-se a necessidade de melhorar a performance do sistema, efetuando-se uma mudança de plataforma. Por esse motivo, todos os registos foram migrados para o novo sistema denominado OASIS/Media Volt. Este trabalho consistiu, também, na reformulação dos tesouros existentes. Denote-se que estes sistemas foram desenhados exclusivamente para as necessidades deste arquivo, o que lhes confere fiabilidade na recuperação de informação. Sublinhe-se que aqui apenas será feita uma análise na ótica do utilizador, sempre com o objetivo de poder definir a política de indexação instituída, não coexistindo o objetivo de um estudo específico ao sistema integrado.

1.2.3.1- Estratégias de busca

a) QUAY

Neste sistema, na interface para o utilizador, estão disponíveis quatro campos de pesquisa:

- 1- FOTOS
- 2- GRÁFICOS

- 3- NOTÍCIAS
- 4- PÁGINAS

Estes documentos são pesquisáveis e recuperáveis consoante a sua especificidade:



1- FOTOS

- data (foto e publicação);
- título;
- autor;
- tema (algumas)

3- NOTÍCIAS

- data;
- título;
- autor;
- tema

2- GRÁFICOS

- data
- título

4- PÁGINAS

- datas

Análise:

Neste sistema, os conteúdos jornalísticos são recuperados pela sua tipologia, sendo esta associada à data de publicação, título, autor e tema. Essencialmente, a pesquisa é assente em quatro grandes parâmetros, ou seja, pelas notícias, páginas na sua totalidade, fotografias e gráficos ou infografias. Alguns documentos estão relacionados por título ou tema. Por exemplo, o gráfico incluso de uma notícia é recuperado com os mesmos descritores dessa mesma notícia. O mesmo acontece com as fotos.

b) OASIS

Neste sistema, na interface para o utilizador, estão disponíveis três campos de pesquisa que se subdividem em outros filtros:

1- TIPO

- fotografia
- crop
- página
- artigo
- artigo texto
- publicação
- edição
- secção
- índice
- destaque

2- DATA DE CONTEÚDO

- arquivo
- modificação

3- TESAUROS

- geográfico
- temático
- personalidades
- fontes
- infraconceito
- locais
- organização
- autores



4- TAGS

- publicado site: sim/não
- texto
- título
- estado

5- DESCRIÇÃO

Análise:

Percebe-se que neste sistema, os filtros de pesquisa são mais vastos, havendo mais pontos de acesso e, conseqüentemente, uma oferta mais específica na recuperação da informação. Há mesmo um maior refinamento daquilo que se procura, nomeadamente na tipologia documental, como se pode perceber com o campo “crop” que se trata de parte de uma fotografia recortada ou, então, com os filtros “artigo” e “artigo texto” em que o primeiro faz uma pesquisa pelos títulos e subtítulos e outros paratextos e o segundo pelo texto na íntegra. Compreende-se que a estrutura dos conteúdos jornalísticos assume uma grande importância, não só pelo seu conjunto mas, também, pelas suas particularidades. Outro aspeto importante no momento da pesquisa será a data em que o documento foi arquivado e/ou modificado. Isto porque, no caso das fotografias, podem ser publicadas mais do que uma vez. O terceiro campo de pesquisa incide essencialmente com o conteúdo dos documentos, coincidindo, na sua maioria, com os tesouros existentes.

2. Tesouros: gestão e classes

2.1- Gestão de construção

Na construção dos tesouros, é utilizado o **método indutivo**, já que vai sendo feita mediante as necessidades, no momento da indexação. É a linguagem natural do jargão jornalístico que serve de base para a construção dos diferentes tesouros. Com base na terminologia linguística publicada, as necessidades de recuperação de informação são correspondidas. Desta forma, percebe-se que no Arquivo e Centro de Documentação do Grupo Controlinveste, os tesouros são de indexação pelo facto de lá serem elaborados de uma forma evolutiva. Tal faz todo o sentido, na medida em que os documentos primários indexados são de proveniência interna e de necessidade de recuperação permanente para a prática jornalística da redação. Sendo assim, o critério usado na admissão de termos dos tesouros é de autoridade ou bibliográfica, já que são oriundos dos textos produzidos na redação para publicação.

2.2- Classes e Tipologia

A linguagem documental adotada cumpre o princípio da pós-coordenação. Desta forma, são vários os Tesouros utilizados no Arquivo e Centro de Documentação do Grupo Controlinveste no sistema de informação QUAY.

- A- AUTORES
- B- EMPRESAS
- C- EQUIPAS DESPORTIVAS
- D- FONTES
- E- GEOGRÁFICO
- F- INSTITUIÇÕES
- G- PERSONALIDADES
- H- SECÇÕES
- I- TIPO
- J- TEMAS

Quanto ao tema, podemos considerar que o tesouro “Temas” é geral e multidisciplinar, pois a cobertura de assuntos abrange variadíssimas áreas do conhecimento. Contudo, os restantes tesouros afunilam para uma especificidade pelo facto de agruparem nomes de pessoas individuais ou coletivas; equipas desportivas; localizações geográficas ou, mesmo, partes constituintes do jornal, como é o caso dos tesouros “Secções” e “Tipo”.



Desta forma, pode-se considerar que o Arquivo e Centro de Documentação do Grupo Controlinveste também recorre a tesouros especializados.

Passemos à análise dos diferentes tesouros do Arquivo e Centro de Documentação do Grupo Controlinveste, analisando os termos quanto à sua relevância e composição, assim como a estrutura e relações básicas de cada tesouro. Contudo, será dada uma maior ênfase ao tesouro “Temas” por se tratar daquele que é mais utilizado na indexação.

2.2.1- Tesouro AUTORES

a) nomes de pessoas individuais

Segundo a NP 4036 (1992), os nomes próprios de pessoas são frequentemente excluídos de um tesouro e quando incluídos, a sua forma deve obedecer a regras de catalogação. Neste caso, justifica-se a existência deste Tesouro tendo em conta a natureza deste centro de documentação, visto interessar a autoria de determinados documentos, sejam as notícias como as fotos. Contudo, não segue, quanto à forma, as regras de catalogação, que determina a forma invertida. Aqui, os nomes próprios de pessoas individuais surgem em forma direta, o que se coaduna com a proximidade à linguagem natural.

Exemplos:

ADÉRITO SERRÃO ADELINO CUNHA ADRIANO MOREIRA

Denote-se que um número significativo de descritores surge com **Notas explicativas (NE)**:

Exemplos:

ALBANO MATOS NE: Jornalista DN

Maria José Rodrigues

ALBERTO JOÃO JARDIM
NE: colaborador de *O Diabo*, presidente
do governo da Região Autónoma da
Madeira

Segundo a NP 4036 (1992), as notas explicativas podem acompanhar os termos para limitar e precisar o sentido em que estes devem ser utilizados e assim excluir outros possíveis significados, acabando, muitas vezes, por servirem de instruções para os indexadores. Na verdade, acabam por ter a função de limitar o âmbito da aplicação do termo. Neste caso, as notas explicativas contextualizam o indexador na utilização dos termos, servindo de notas biográficas do autor, principalmente relacionadas com a sua ligação profissional.

Alguns termos surgem como não preferenciais, como é o caso de nomes completos ou menos conhecidos.

Exemplo:

ALBERT GEA UF: Albert Julian Gea
NE: Reuters

No entanto, o não-descritor está ausente na lista alfabética, mas recuperável pois surge como sinónimo do descritor, assim como é indicado na NP 4036 (1992), já que define **não-descritor** como sendo o sinónimo ou quasi-sinónimo de um descritor. Não pode ser atribuído a documentos, mas serve de entrada num tesouro, sendo remetido através de uma nota (USE) para o descritor apropriado.

b) Nomes de entidades, instituições ou pessoas coletivas

Neste tesouro, surgem algumas instituições surgem como autoras, nomeadamente de fotografias. Terminologicamente, surgem na forma direta, tal como surge na linguagem natural. No caso de existir duas grafias, o termo preferencial ou descritor será sempre aquele mais utilizado na linguagem natural.

Exemplos:

ACORMEDIA UF: Açormedia
ADIDAS



2.2.2- Tesouro EMPRESAS

Este tesouro apresenta termos representativos de nomes de empresas, portuguesas e estrangeiras que possam constituir assunto. Como é bastante frequente nestes casos, várias empresas estão associadas a acrónimos, surgindo a necessidade de eleger o termo preferencial. Percebe-se que não existe uma regra uniforme quanto à forma, continuando o critério de se considerar descritor aquele que é mais utilizado e pela qual a empresa é mais conhecida. A NP 4036 (1992: 11) indica que “os acrónimos não devem ser utilizados como descritores, excepto se estiverem largamente divulgados (...) e a forma por extenso for raramente utilizada ou mesmo desconhecida”.

Exemplos:

AGENCE FRANCE PRESSE UF: AFP
AGERE UF: Empresa de Águas Fluentes e Resíduos de Braga

2.2.3- Tesouro EQUIPAS DESPORTIVAS

Neste tesouro, das equipas desportivas constam clubes de futebol, seleções e associações desportivas. Também surgem não descritores ou, então, notas explicativas como nos tesouros anteriormente referidos.

Exemplos:

ASSOCIAÇÃO DESPORTIVA OS LIMIANOS UF: Limianos
ELÉCTRICO NE: Clube desportivo

2.2.4- Tesouro FONTES

Trata-se de um tesouro referente a jornais, agências noticiosas ou revistas, como fontes do documento indexado. Muitos descritores surgem hierarquizados numa relação partitiva, no caso de publicações dependentes ou incluídos noutros, como é o caso de suplementos ou revistas de jornais.



Exemplos:

<p>TG 24 HORAS TE BITS & BYTES UF: Bits TE EM CASA</p> <p>TG CORREIO DA MANHÃ TE CORREIO DA MANHÃ REVISTA TV TE MAGAZINE DOMINGO</p> <p>TG DIÁRIO DE NOTÍCIAS UF: DN TE DIÁRIO DE NOTÍCIAS ECONOMIA TE1 DN ECONOMIA SUPLEMENTOS TE DNA TE DN AMBIENTE</p>

2.2.5- Tesouro GEOGRÁFICO

Neste tesouro, estão incluídos países e suas capitais e cidades, distritos, assim como penínsulas, dispostos hierarquicamente numa relação partitiva. Percebe-se nesta estrutura que a sua construção seguiu o método indutivo já que nem todos os países alinham com as suas capitais como termos específicos, por exemplo. Estes termos possuem duas modalidades de aplicação, isto é, podem ser aplicados quando constitui assunto básico ou, então, quando constitui aspeto geográfico associado a qualquer tema.

Exemplos:

<p>TG ALEMANHA TE MUNIQUE</p> <p>TG BRASIL TE RIO DE JANEIRO</p> <p>TG BULGÁRIA TE SOFIA</p> <p>TG CABO VERDE TE ILHA DO SAL</p> <p>TG ESTADOS UNIDOS DA AMÉRICA UF: EUA TE FLÓRIDA</p> <p>TE NOVA IORQUE</p> <p>TG PENÍNSULA IBÉRICA TE ESPANHA TE PORTUGAL TE1 AÇORES TE1 DISTRITO DA GUARDA TE2 AGUIAR DA BEIRA</p>



2.2.6- Tesouro INSTITUIÇÕES

Trata-se de um tesouro onde constam nomes de instituições e associações, que não constituem empresas. Alguns descritores têm qualificadores, no caso de homonímia. Também surgem não descritores ou, então, notas explicativas. Não existe uma uniformidade na escolha dos acrónimos ou da sua forma por extenso como descritores, continuando o critério de se considerar descritor aquele que é mais utilizado e pela qual a instituição é mais conhecida.

Exemplo:

AEP [ASSOCIAÇÃO EMPRESARIAL DE PORTUGAL] NE: Associação Empresarial de Portugal
AMI UF: Assistência Médica Internacional
ASSOCIAÇÃO COMERCIAL DO PORTO UF: ACP [Associação Comercial do Porto]

2.2.7- Tesouro PERSONALIDADES

Neste tesouro, estão presentes nomes próprios individuais que não são autores dos documentos primários produzidos na redação, mas sim determinadas personalidades da vida política, social, entre outras. Também não seguem, quanto à forma, as regras de catalogação, que determinam a apresentação pela forma invertida. Tal se coaduna com a proximidade com a linguagem natural e os termos utilizados no momento da pesquisa.

Alguns descritores têm qualificadores, aqui entre parêntesis retos, no caso de homonímia. Também surgem não descritores sempre que se justifique a inclusão de outra forma do nome também conhecido ou, então, notas explicativas como forma de introduzir informações acerca dessa mesma pessoa.

Exemplos:

ABEL JORGE PEREIRA DA SILVA UF: Abel Silva
ABÍLIO SILVA [comerciante]

Maria José Rodrigues

ABÍLIO SILVA [deputado municipal]
UF: Abílio Sousa e Silva
NE: Deputado municipal, PSD de
Caminha

2.2.8- Tesouro SECÇÕES

Este tesouro apresenta termos relativos às secções ou partes dos jornais, revistas ou suplementos, indicando, de alguma forma, o assunto principal a que está associado o documento a indexar. Na verdade, implica a estrutura intelectual onde o documento a indexar se inclui dentro da publicação, dando algumas informações ao indexador durante a primeira fase da indexação. Muitos descritores surgem hierarquizados, visto se tratar de partes de uma publicação. Alguns apresentam notas explicativas e, também, não descritores.

Exemplos:

TG DESPORTO
TE DNSport
TE MUNDIAL 2006
TG DIA DAS COMADRES
NE: secção do jornal Açoriano Oriental
TG ECONOMIA E NEGÓCIO UF:
Economia & Negócios

2.2.8- Tesouro TIPO

Este tesouro apresenta termos relativos à tipologia de documento e/ou abordagem, que servem para auxiliar a pesquisa e são considerados *infra conceitos*. Na verdade, este tesouro constitui um aspeto diferenciador e específico para os conteúdos jornalísticos. Sem dúvida que a tipologia do texto constitui um ponto de acesso importante no momento da pesquisa, visto que o assunto está organizado de uma forma específica, seja ela ao nível do conteúdo ou da forma. Alguns não-descritores apresentam a forma de abreviatura, visto ser a forma também conhecida no meio jornalístico.

Exemplos:

GRÁFICO
ÍCONE
INQUÉRITO UF: INQ
REPORTAGEM UF: REP
FOTOLEGENDA UF: FL
ENTREVISTA UF: ENT
CRÍTICA



CRÓNICA
EDITORIAL
SONDAGEM

2.2.9- Tesouro TEMAS

Trata-se do tesouro com uma estrutura mais complexa, de temática geral e multidisciplinar, cuja apresentação é hierarquizada e alfabética. A hierarquia é pluridimensional, visto existirem diversos níveis.

Exemplo:

TG MEIOS DE PRODUÇÃO AGRÍCOLA
TG1 GADO
TE BOVINO
TE1 BOI
TE1 NOVILHO
TE1 TOURO
TE1 VACA
TE2 VACA LEITEIRA
TE2 VACA REPRODUTORA
TE1 VITELA
TE CABEÇA DE GADO
TE CAPRINO
TE CERVÍDEOS
TE EQUINO
TE OVINO
TE SUINO

Analisemos, em primeiro lugar, os termos de topo, a partir dos quais são desdobrados termos específicos numa estrutura arborescente ou hierarquizada.

TERMOS DE TOPO:

ABASTECIMENTO DE GÁS
ACIDENTES
ACTIVIDADE ECONÓMICA
AGRICULTURA E SILVICULTURA UF:AGRICULTURA
AGRO ALIMENTAR

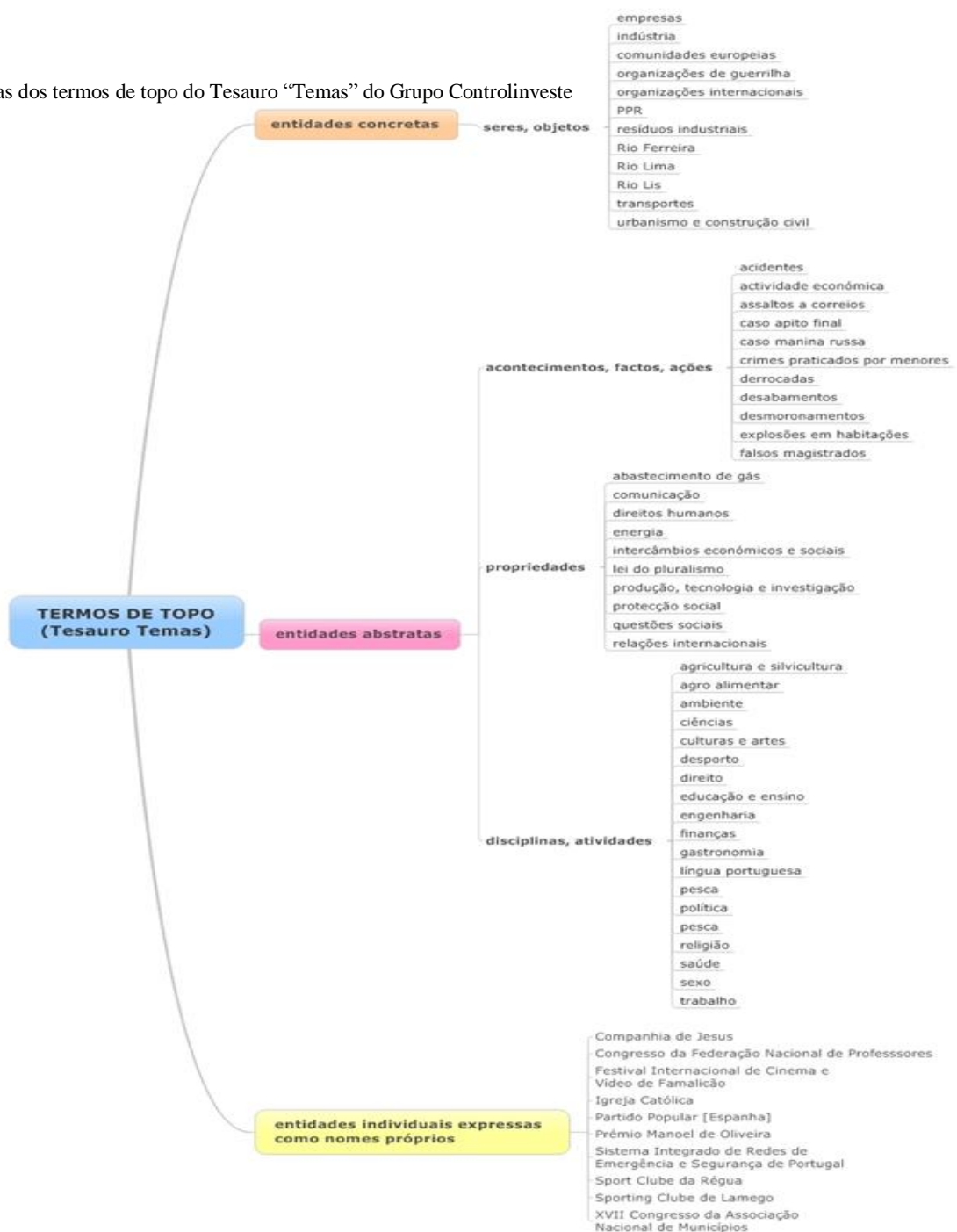


AMBIENTE
ASSALTOS A CORREIOS UF:ASSALTOS A CTT
CASO APITO FINAL
CASO MENINA RUSSA
CIÊNCIAS
COMPANHIA DE JESUS
COMUNICAÇÃO
COMUNIDADES EUROPEIAS
CONGRESSO DA FEDERAÇÃO NACIONAL DE PROFESSORES
CRIMES PRATICADOS POR MENORES
CULTURA E ARTES UF:ARTE E CULTURA
DERROCADAS
DESABAMENTOS UF:DESLIZAMENTOS
DESMORONAMENTOS
DESPORTO
DIREITO
DIREITOS HUMANOS
EDUCAÇÃO E ENSINO
EMPRESAS
ENERGIA
ENGENHARIA
EXPLOSÕES EM HABITAÇÕES
FALSOS MAGISTRADOS
FESTIVAL INTERNACIONAL DE CINEMA E VIDEO DA SERRA DA ESTRELA UF:CINEECO
FESTIVAL INTERNACIONAL DE CINEMA E VIDEO DE FAMALICÃO UF:FAMAFEST
FINANÇAS
GASTRONOMIA
IGREJA CATÓLICA
INDÚSTRIA
INTERCÂMBIOS ECONÓMICOS E SOCIAIS
LEI DO PLURALISMO
LINGUA PORTUGUESA
ORGANIZAÇÕES DE GUERRILHA
ORGANIZAÇÕES INTERNACIONAIS
PARTIDO POPULAR [ESPANHA]
PESCA UF:PESCAS
POLÍTICA
PPR UF:PLANOS POUPANÇA REFORMA
PRÉMIO MANOEL DE OLVEIRA
PRODUÇÃO, TECNOLOGIA E INVESTIGAÇÃO
PROTECÇÃO SOCIAL
QUESTÕES SOCIAIS
RELAÇÕES INTERNACIONAIS
RELIGIÃO
RESÍDUOS INDUSTRIAIS
RIO FERREIRA
RIO LIMA
RIO LIS
SAÚDE
SEXO
SISTEMA INTEGRADO DE REDES DE EMERGÊNCIA E SEGURANÇA DE PORTUGAL
UF:SIRESP
SPORT CLUBE DA RÉGUA
SPORTING CLUBE DE LAMEGO
TRABALHO
TRANSPORTES
URBANISMO E CONSTRUÇÃO CIVIL
XVII CONGRESSO DA ASSOCIAÇÃO NACIONAL DE MUNICÍPIOS

Análise:

É perceptível que estes termos são fruto do método indutivo, já antes referido, pois trata-se de descritores relativos a casos aleatórios que foram surgindo como notícia ao longo de várias publicações. Não existe propriamente uma uniformidade de critérios quanto à forma dos termos, percebendo-se uma abordagem de procedimentos de construção analítica, pouco crítica, mas apenas tendo em atenção às reais necessidades do público-alvo. Contudo, embora de uma forma genérica, já é possível delinear o perfil temático dos conteúdos existentes no arquivo e centro de documentação do Grupo Controlinveste. Categorizemos em grupos, as noções representadas por estes termos de topo:

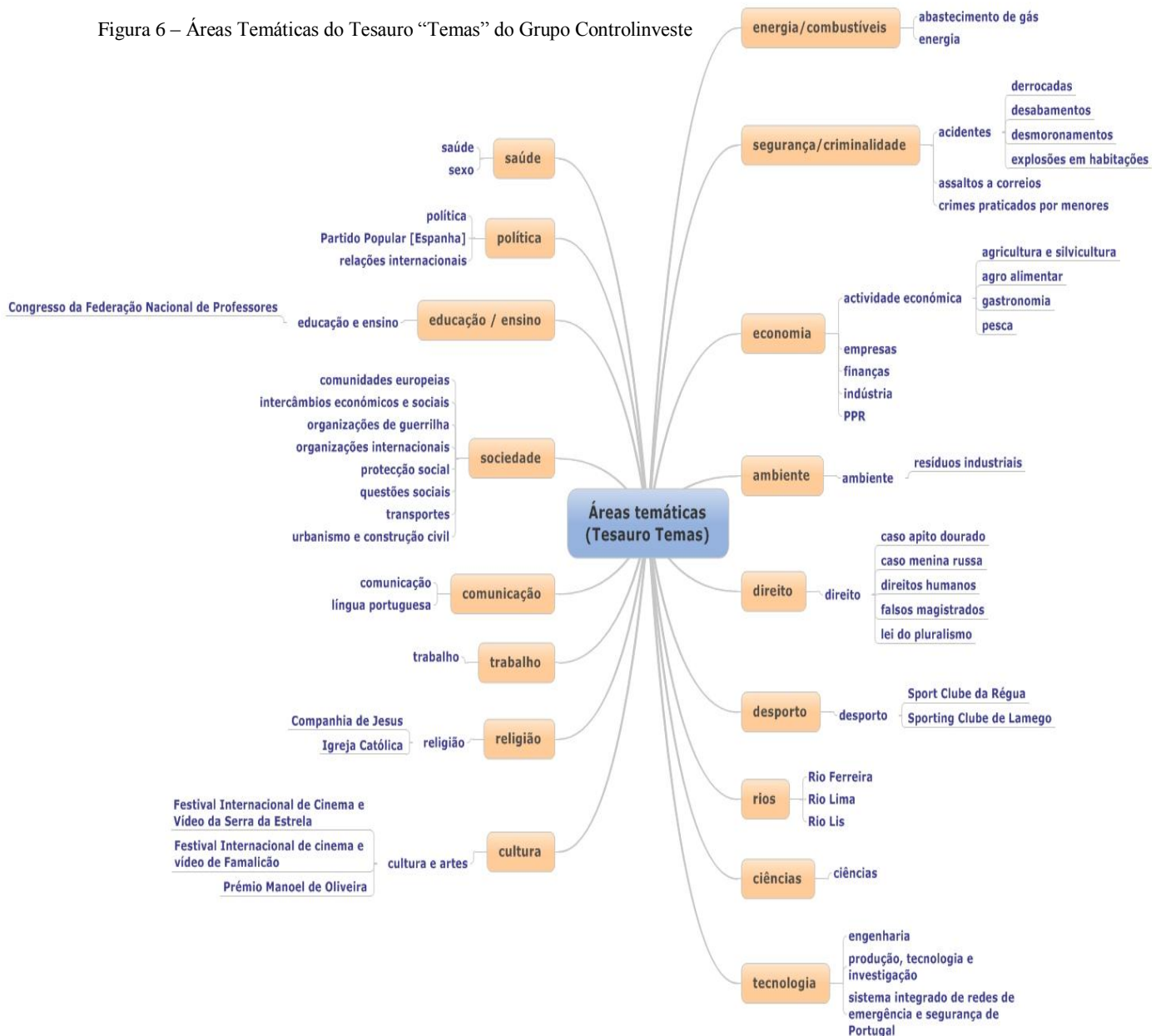
Figura 5 – Categorias dos termos de topo do Tesouro “Temas” do Grupo Controlinveste



Depreende-se deste esquema que vários termos de topo deveriam estar dependentes de outros, tais como os nomes de rios, que deveriam ser integrados no Tesouro geográfico, e também os nomes individuais expressos como nomes próprios. Percebe-se, também, a existência de sinónimos alinhados ao mesmo nível, que é o caso dos termos “derrocadas”, “desabamentos” e “desmoronamentos”, o que confirma a existência de casos aleatórios antes referidos.

Contudo, é de salientar a diversidade temática presente, o que se justifica visto se tratar de documentos oriundos de títulos da imprensa generalista, como é o caso do Grupo Controlinveste. Vejamos agora, as diferentes áreas temáticas que abrangem este tesouro “Temas” a partir dos seus termos de topo (em azul):

Figura 6 – Áreas Temáticas do Tesouro “Temas” do Grupo Controlinveste



Maria José Rodrigues

É notório nesta organização temática a presença das várias áreas que os jornais diários abordam nas suas publicações, coincidentes, na sua maioria, com as diversas secções. Observemos a estrutura disponível nas páginas web do “Jornal de Notícias” (www.jn.pt) e do “Diário de Notícias” (www.dn.pt) que, sendo os jornais mais generalistas, servem aqui de exemplo para este aspeto, estrutura essa equivalente à versão em papel:



Ilustração 5 – Menus das páginas em linha do “Jornal de Notícias” e “Diário de Notícias”

É perceptível a coincidência das áreas temáticas do tesouro “Temas”, antes esquematizadas, com as diversas secções dos jornais em questão, cujos conteúdos são indexados pelo mesmo tesouro. Não esquecendo que os tesouros do Arquivo e Centro de Documentação do Grupo Controlinveste são construídos pelo método indutivo, claramente se entende esta coincidência. Vejamos então as áreas comuns:

Jornal/versão	Jornal de Notícias	versão	Diário de Notícias	versão	Áreas temáticas do Tesouro ‘Temas’
secções	Política	papel/digital	Política	papel/digital	Política
	Sociedade	papel/digital	Sociedade	papel	Sociedade
	Segurança	papel/digital	Segurança	papel	Segurança/Criminalidade
	Economia	papel/digital	Economia	digital	Economia
			Bolsa	papel	
	País	digital	Portugal	digital	
	Norte-Sul	papel	Cidades	papel	
	Mundo	papel/digital	Globo	papel/digital	
	Desporto	papel/digital	Desporto	papel/digital	Desporto
	Cultura	digital	Cartaz	digital	Cultura
			Guia	papel	
	Artes/Vidas	papel	Artes	papel/digital	
			TV e Media	digital	
Gente	digital	Pessoas	digital		
Tecnologia	papel/digital	Ciência	papel	Ciências	
				Tecnologia	

Tabela 5 – Análise comparativa das secções dos jornais com as Áreas temáticas do Tesouro Temas



Claro que as restantes áreas temáticas são perfeitamente incluíveis nas diversas secções, principalmente aquelas que abarcam temas mais abrangentes, tais como “Sociedade”, “Política” e “Economia”.

2.3- Categorias temáticas, termos, estrutura e relações básicas

Considerando o Tesouro “Temas” aquele com mais riqueza estrutural e conteudal para uma análise mais pormenorizada, foram escolhidos três termos de topo para proceder a diferentes estudos às suas estruturas arborescentes, no que diz respeito às categorias temáticas, à estrutura e relações básicas assim como ao tipo de termos. Desta forma, far-se-á as respetivas explorações aos termos de topo “**Comunicação**”; “**Actividade económica**” e “**Trabalho**”, considerando as suas estruturas arborescentes como se de microtesauros se tratassem, conforme são apresentados em anexo (ANEXO 2).

2.3.1- Categorias temáticas no Microtesauro COMUNICAÇÃO

Aqui far-se-á uma análise às categorias temáticas da estrutura arborescente do termo de topo “Comunicação”. O objetivo é perceber quais foram as áreas de conhecimento consideradas na construção e organização desta estrutura, na perspectiva de que os conteúdos jornalísticos deste arquivo abrangem uma vasta área informativa, pois é a partir deles que essa construção se efetua através do método indutivo. Foi escolhido este microtesauro, para ser efetuada esta análise, devido ao facto de mais se aproximar ao conteúdo já explorado na primeira parte desta tese. Assim, é mais fácil avaliar a pertinência das áreas temáticas incluídas nesta estrutura. Começamos por fazer uma breve abordagem ao termos genéricos que desta árvore fazem parte.

TT COMUNICAÇÃO

TG COMUNICAÇÃO EM GRUPO

TG DOCUMENTAÇÃO UF: Documentação científica, documentação técnica

TG DOCUMENTO

TG INDÚSTRIA DA COMUNICAÇÃO

TG INFORMAÇÃO E TRATAMENTO DA INFORMAÇÃO

TG INFORMATICA E PROCESSAMENTO DE DADOS

TG MATERIAL AUDIOVISUAL

TG MEIOS DE COMUNICAÇÃO

TG POLÍTICA DA COMUNICAÇÃO UF: Desenvolvimento das comunicações

TG SISTEMA DE COMUNICAÇÃO

TG SISTEMA DOCUMENTAL

Dos onze termos genéricos, apreende-se a inclusão de três grandes classes informacionais: **informação**, **documentação** e **comunicação**, tendo cada uma subclasses, que apresentam diferentes aspetos de abordagem, diretamente relacionadas com os termos genéricos (em azul), assim como o que o esquema explicita:

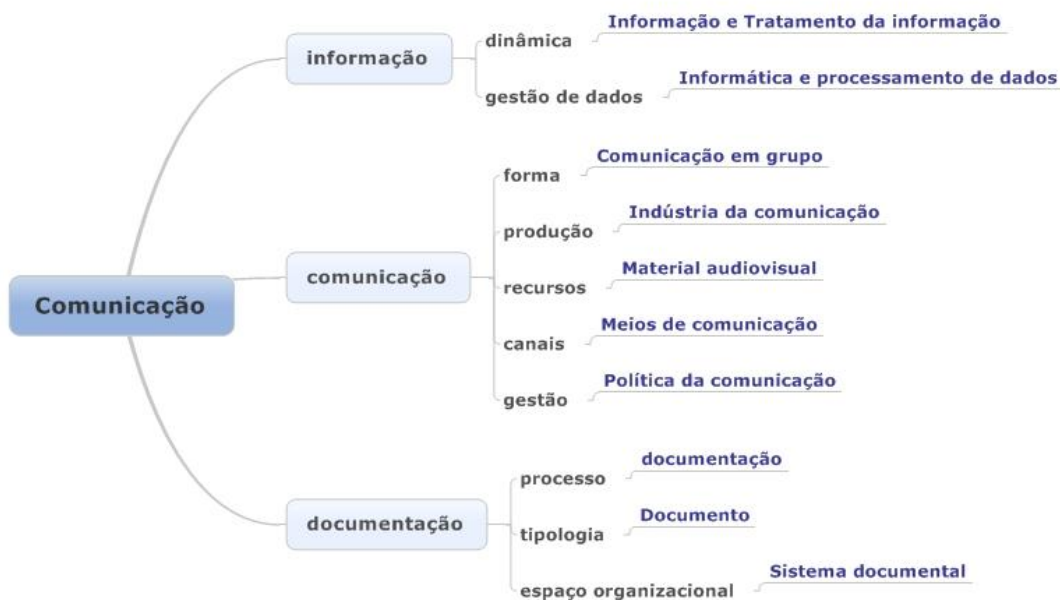


Figura 7 – Categorias temáticas do Microtesauro “Comunicação”

Esta estrutura tripartida revela a coerência conteudal deste microtesauro, no sentido que, ao retirar dele as grandes unidades informacionais, conseguimos entender o que pode representar ao nível do conhecimento, que está incluído nos diversos conteúdos jornalísticos. Para além disso, compreende-se que, no momento da indexação, vários aspetos sejam abordados dentro de uma grande classe que, aqui, está representada pelo termo de topo “Comunicação”. Se antes foi afirmado que o tesouro é um instrumento de representação da informação e do conhecimento, mais se entende ao analisá-lo desta forma. Aliás, pode ser feita uma leitura teórica sobre comunicação através deste esquema, entendendo que o processo comunicacional engloba a **informação** nas suas diferentes vertentes, desde a origem, passando pela política de **acesso** até ao seu **tratamento**, tarefas do **profissional da informação**. Nesta dimensão, urge uma gestão de dados informacionais através de um **sistema informático** com o intuito de também estabelecer a comunicação dessa mesma informação através do **profissional da informática**. Tal será produzida pela **indústria da comunicação** em diversas plataformas pelo **profissional da comunicação**, com recurso a **material audiovisual**, de forma a difundir a informação produzida nos diferentes **meios de comunicação**

social, que servem de canais comunicativos, tais como a **imprensa**, a **radiodifusão** ou a **televisão**. Contudo, a comunicação tem de ser regida por uma política de gestão. Claro está que não é possível abordar a dimensão comunicacional sem recorrer ao suporte informacional que é o **documento**. Este, em diferentes tipologias, sejam elas de ordem física ou intelectual, sofre processos, tais como **aquisição**, **armazenagem**, **gestão** e **difusão**, de forma a servir as dinâmicas informacionais e comunicacionais.

Esta abordagem, embora redutora e que apenas serve de amostra, permite compreender que existe um cuidado em estruturar intelectualmente os tesouros no Arquivo e Centro de Documentação do Grupo Controlinveste, visto que as palavras em destaque são descritores, ou parte deles, retirados da estrutura em análise.

2.3.2- Termos, estrutura e relações básicas

Após uma análise intelectual do microtesouro “Comunicação” far-se-á uma abordagem mais quantitativa aos termos e suas relações básicas nas três estruturas hierarquizadas “Actividade económica”; “Comunicação” e “Trabalho”.

A análise assenta em quatro pontos:

a) Levantamento do número total de descritores e não-descritores

Objetivos: quantificar a amostra e a qualificar a dimensão da estrutura; quantificar os termos quanto à sua relevância como termos preferenciais e não-preferenciais.

Microtesouro	Nº descritores	Nº não-descritores	TOTAL TERMOS (descritores e não- descritores)
Actividade económica	132	73	205
Comunicação	240	146	386
Trabalho	188	141	329
TOTAL	560	360	920

Tabela 6 – Descritores e Não-descritores dos microtesouros “Actividade Económica”, “Comunicação” e “Trabalho”

O universo de descritores dos três microtesouros é de 560. Depreende-se que têm uma dimensão considerável, visto todos ultrapassarem a centena de termos preferenciais a

Maria José Rodrigues

serem utilizados na indexação de conteúdos jornalísticos com estas áreas temáticas. O número total de não-descritores é de 360, o que constitui 39,1% da totalidade dos termos existentes. Ora, tal confere que cerca de 60,9% dos termos destes três microtesauros são descritores, o que revela que a maioria são termos preferenciais a utilizar na indexação e recuperação dos conteúdos jornalísticos. No entanto, o número de não-descritores é, de facto, considerável, o que pressupõe a existência de controlo de vocabulário na construção dos tesauros. Tal obriga uma outra análise, a de perceber quais são os critérios para a eleição dos não-descritores que fazem-nos distinguir dos termos preferenciais e, por isso, constituir uma reciprocidade.

b) Levantamento do número de descritores acompanhados de não-descritores

Objetivo: avaliar o grau de controlo do vocabulário na construção do tesauro e a relação de equivalência que apresenta.

Microtesauro	Nº descritores c/ não-descritores
Actividade económica	44
Comunicação	83
Trabalho	76
TOTAL	203

Tabela 7 – Nº descritores com não-descritores dos microtesauros “Actividade Económica”, “Comunicação” e “Trabalho”

Aqui percebemos que dos 560 descritores, cerca de 36,25% é acompanhado de não-descritores, ou seja, menos de metade. No entanto, tratando-se de um número significativo, fez-se um levantamento da natureza dos não-descritores que levaram à sua eleição como termos não-preferenciais, que poderão ajudar a delinear os objetivos desta eleição na política de indexação do Arquivo e Centro de documentação do Grupo Controlinveste.

Natureza dos não-descritores exemplos

Sinónimos

ANÁLISE ECONÓMICA UF avaliação económica
RIQUEZA UF fortuna
FORMULÁRIO UF impresso
RECRUTAMENTO UF admissão do pessoal
REFORMADO UF aposentado



Natureza dos não-descritores	exemplos
Plural	SONDAGEM UF sondagens CIRCUITO TURÍSTICO UF circuitos turísticos LIVRO UF livros JORNALISTA UF jornalistas
Singular	DESPEDIMENTO UF despedimentos
Acrónimos	SOTERRADOS UF soterrado PRODUTO INTERNO BRUTO UF PIB
Desenvolvimento de acrónimos	TIM UF Programa Integrado Mediterrânico GED UF Gestão Electrónica de Dados OCR UF Leitura Óptica de Caracteres
Termos relacionados	DICIONÁRIO UF léxico INDÚSTRIA DO LIVRO UF encadernação PROFISSIONAL DA INFORMAÇÃO UF bibliotecário, arquivista, documentalista TRATAMENTO DA INFORMAÇÃO UF tratamento computadorizado
Outra grafia / outra língua	TESAURO UF thesaurus EQUIPAMENTO INFORMÁTICO UF hardware INTERNET UF web
Tradução para português	SOFTWARE UF programa de computador
Abreviatura	TELEVISÃO DIGITAL UF TV digital

Tabela 8 – Natureza dos Não-descritores dos microtesauros “Actividade Económica”, “Comunicação” e “Trabalho”

Nota: os mais usuais são os casos de sinonímia e os dos conceitos relacionados. A opção por estes últimos aponta para a necessidade de diminuir a carga lexical do tesouro no momento da indexação, ocupando o espaço do vazio que existe em relação à inexistência de relações associativas. Na verdade, estes não-descritores funcionam como termos relacionados, visto serem recuperáveis no momento da pesquisa, apesar de não serem aplicados durante a indexação. O controlo da sinonímia também indica o cuidado que existe em construir um vocabulário consistente para que a indexação e a recuperação dos conteúdos jornalísticos também o sejam. Contudo, nota-se que não existe uma uniformidade nas opções, percebendo-se que subsiste uma maior preocupação em estabelecer uma aproximação à linguagem natural.

c) Levantamento dos descritores em forma de termos simples e termos compostos

Objetivo: avaliar a composição dos termos e a sua relação com a linguagem natural e a especificidade na representação dos assuntos.

Agora, para avaliar a composição dos termos e a sua relação com a linguagem natural e a especificidade na representação dos assuntos, far-se-á o levantamento dos descritores

em forma de termos simples e termos compostos nos microtesauros “Actividade económica”; “Comunicação” e “Trabalho”.

Microtesauro	Actividade económica	Comunicação	Trabalho	TOTAL
Nº termos simples	19	78	22	119
Nº termos compostos	113	162	166	441
expressões adjetivas	66	74	66	206
expressões prepositivas	38	76	83	197
outros	9	12	17	38
TOTAL	132	240	188	560

Tabela 9 – Termos simples e termos compostos dos microtesauros “Actividade Económica”, “Comunicação” e “Trabalho”

Como se percebe, a grande maioria dos termos tem a forma composta, mais precisamente 78,75% da totalidade dos três microtesauros, sendo 21,25% termos simples. Entre os termos compostos, nota-se uma preferência pelas expressões adjetivas (46,7%), embora com uma margem muito pequena em relação às expressões prepositivas (44,7%), o que se considera, mediante esta amostra, que serão equivalentes. Os outros termos compostos são constituídos por expressões nominais e prepositivas em simultâneo, unidas com uma conjunção coordenativa copulativa, como é exemplo “Administração e remuneração do pessoal”; por dois nomes seguidos, como é o caso “Sítio internet”; por expressões prepositivas e adjetivas em simultâneo, como serve de exemplo “Estratégia europeia de emprego”; ou também por expressões adverbiais “Actividade não assalariada”, entre outros. É claro que estes casos deverão ser revistos para que sejam substituídos por formas adjetivas ou prepositivas, que serão as mais indicadas, segundo a NP 4036 (1992). Porém, constituem apenas 8,6% da totalidade dos termos compostos das três estruturas aqui analisadas, o que não confere grande significância. Percebe-se, porém, que com estes termos compostos existe uma necessidade de se aproximar da linguagem natural assim como uma preocupação em diminuir o número de termos por documento a indexar, fazendo com que haja uma maior especificidade nos assuntos a traduzir.

d) Levantamento do número de termos quanto à sua relação de subordinação (termos gerais e específicos)

Objetivo: compreender a relação hierárquica entre os termos e quantificar cada nível.

Maria José Rodrigues

Por fim, para compreender a relação hierárquica entre os termos e quantificar cada nível, é pertinente fazer um levantamento de termos quanto à sua relação de subordinação (termos gerais e específicos), visto se tratarem de estruturas hierárquicas. Nos três microtesauros, foram identificados sete níveis hierárquicos, ficando identificados e estruturados da seguinte forma:

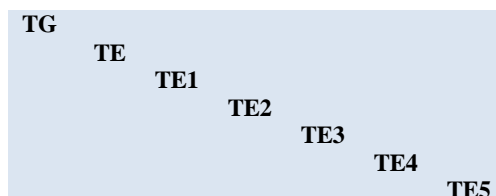


Tabela 10 – Níveis hierárquicos dos termos dos microtesauros “Actividade Económica”, “Comunicação” e “Trabalho”

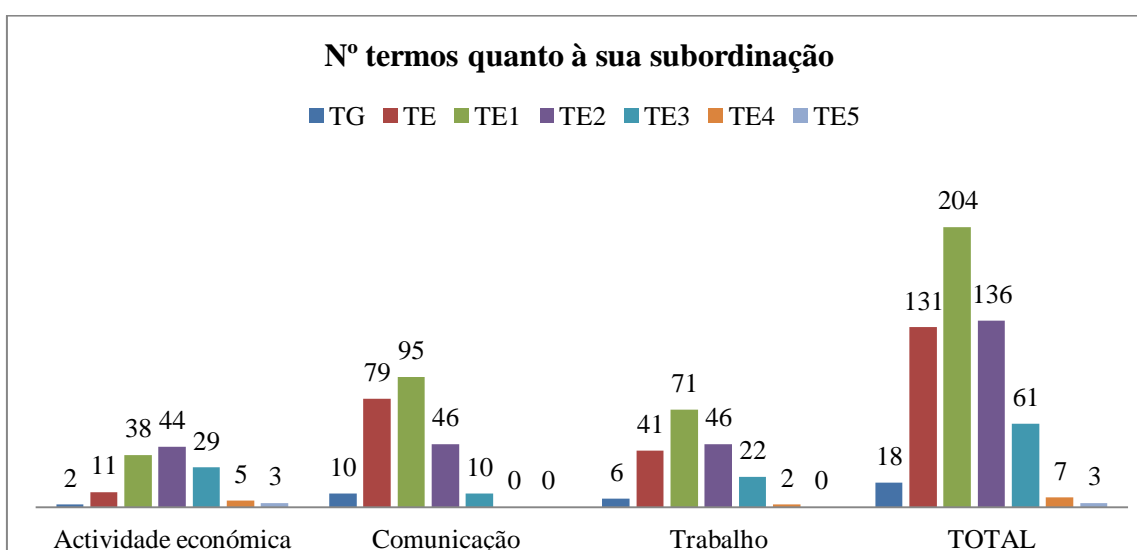


Gráfico 1 – Nº termos quanto à sua subordinação dos microtesauros “Actividade Económica”, “Comunicação” e “Trabalho”

A partir destes gráficos, é perceptível a tendência para uma incidência nos Termos específicos dos terceiros e quartos níveis de hierarquia (TE1 e TE2), embora o segundo nível se encontra muito próximo destes (TE). Ora, daqui depreende-se que o vocabulário integrante dos tesauros do Arquivo e Centro de Documentação tende a ser específico de forma a representar assuntos da mesma ordem. Dessa forma, e segundo a NP 4036 (1998), existem três tipos de relações hierárquicas: genérica; partitiva e de instância. A primeira “identifica a ligação entre uma classe ou categoria e os seus membros ou espécies” (NP 4036, 1998:25), a segunda “a parte está implícita no todo” (NP 4036, 1998:26) e, por fim, a terceira “identifica uma categoria geral (...) em espécime individual” (NP 4036, 1998:27). Nesta perspetiva e analisando as relações de subordinação destes microtesauros, percebe-se que as relações dominantes são a genérica e a de instância, denotando-se que quanto mais se afunila o termo (TE3, TE4 e

TE5) mais este representa um espécime individual. Desta forma, como a maioria dos termos balizam os níveis TE, TE1 e TE2, as relações hierárquicas serão, na sua maioria, genéricas, daí a maioria dos termos representarem membros dos termos genéricos. Já a relação partitiva é escassa nestes três exemplos. Então vejamos três exemplos de hierarquias que comprovam este tipo de relações:

ACTIVIDADE ECONÓMICA	COMUNICAÇÃO	TRABALHO
TE1 POLÍTICA ECONÓMICA	TE1 TELEVISÃO	TE1 RELAÇÕES DE TRABALHO
TE2 POLÍTICA DE DESENVOLVIMENTO	TE2 ESTAÇÃO DE TELEVISÃO	TE2 CONFLITO LABORAL
TE3 DESENVOLVIMENTO SUSTENTÁVEL	TE3 RTP	TE3 SINDICATO
TE4 CIMEIRA DA TERRA		TE4 CONFEDERAÇÃO SINDICAL
TE5 AGENDA 21		

Tabela 11 – Relações hierárquicas genéricas e partitivas nos microtesauros “Actividade Económica”, “Comunicação” e “Trabalho”

Nota: Quanto mais especificidade, mais se aproxima de um exemplo concreto. Assim, os TE1 e TE2 e TE3 são membros ou espécies, enquanto que os TE4 e TE5 já representam espécimes.



3. Política de indexação dos conteúdos jornalísticos do arquivo digital do Grupo Controlinveste

Após traçar o perfil dos instrumentos de indexação que os tesouros constituem, perfila-se que há um objetivo comum, tanto no momento da indexação como no da recuperação da informação: o de precisão na representação dos assuntos dos conteúdos jornalísticos, através de termos mais específicos e compostos com proximidade à linguagem natural, embora exista um nível considerável de controlo de vocabulário, como se constata na existência dos não-descritores.

Agora, há-que analisar a indexação desses mesmos conteúdos, recuperados a partir de pesquisas feitas com alguns termos inclusos nos tesouros do Arquivo e Centro de Documentação do Grupo Controlinveste. Tal tem como objetivo definir as linhas de atuação no momento da indexação e, conseqüentemente, perceber as prioridades na recuperação da informação.

Desta forma, vários documentos recuperados no sistema integrado QUAY do Arquivo e Centro de Documentação do Grupo Controlinveste serão perfilados como exemplos a analisar.

3.1- Documentos da imprensa

Começamos pelos documentos escritos da imprensa, aos quais são feitas questões que serão a orientação desse estudo.

Apesar destes exemplos serem escassos, tendo em conta a dimensão dos conteúdos existentes, consegue-se depreender uma linha de atuação das práticas de indexação aplicadas nestes serviços.

-
- a) **São utilizados descritores com nível específico de subordinação na relação hierárquica do tesouro?**
 - b) **São aplicados mais do que um descritor por documento?**
 - c) **São utilizados os diferentes campos de indexação referentes aos diversos tesouros?**
 - d) **A estrutura do documento (títulos, legendas, lead) serviu de fonte para a indexação?**
 - e) **Os descritores abrangem o/os assunto/os dominante(s) do documento?**
-



-
- f) A indexação responde às questões essenciais de um texto jornalístico: *o quê?; quem?; onde?; quando?; como?; porquê?*
- g) Percebe-se a relação entre a indexação e a secção do jornal onde o documento está publicado?
- h) A tipologia documental distingue-se na indexação com a aplicação de infra-conceitos?
- i) A política editorial está implícita na indexação?
- j) Depreende-se a presença da linguagem natural nos termos controlados utilizados na indexação?
- k) Percebe-se a coexistência das indexações automática, manual ou semiautomática?
- l) A indexação confere um grau de especificidade ou exaustividade?
-



Exemplo 1:

FICHA:
Proprietário: JN
Estado: Arquivado
Data da Notícia: 08/01/2009

CAMPOS DE INDEXAÇÃO
Temas:
RECESSÃO ECONOMICA
FALÊNCIA
DESPEDIMENTO COLECTIVO
Instituições:
AUTORIDADE PARA AS CONDIÇÕES
DE TRABALHO

Dados de Publicação
Publicação: JORNAL DE NOTÍCIAS
Edição: PORTO
Secção: PRIMEIRO PLANO

Página: 2

Análise:

Foi feita uma pesquisa com o termo composto (expressão adjetival) “RECESSÃO ECONÓMICA”. Trata-se de um termo específico de nível 1, subordinado ao termo específico “Ciclo económico” e ao termo genérico “Crescimento económico”. Esta estrutura implica a indexação de documentos que abordem a dinâmica económica, no seu ciclo de crescimento e recessão. Foram aplicados três termos do tesouro Temas “RECESSÃO ECONÓMICA”; “FALÊNCIA” e “DESPEDIMENTO COLECTIVO” e um do de Instituições “AUTORIDADE PARA AS CONDIÇÕES DE TRABALHO”.

Este termo está presente no título, através do seu acrónimo; no subtítulo e no *lead*. O segundo encontra-se igualmente no *lead*. O primeiro é o assunto genérico da notícia, relacionado com o antetítulo “*As consequências da crise económica*”. Sendo **Crise económica** um não-descritor ou descritor não preferencial, sinónimo de **Recessão económica**, o tema central da notícia, abordada numa faceta consequente, é recuperável. Daqui se percebe a proximidade dos termos controlados com a linguagem natural presente no texto da notícia, o que parece estar implícita uma indexação semiautomática. Desta forma, a indexação realizada responderá às questões *o quê?*; *quem?*; *como?* e *porquê?*, tendo sido utilizados dois dos tesouros existentes: temas e instituições. Para além do conteúdo do texto, é perceptível a utilização dos títulos e do *lead* para a eleição dos respetivos descritores. Esta notícia pode ser recuperada numa pesquisa sobre crise e/ou recessão económica. Para além disso, é pesquisável para quem pretender recuperar documentos relativos à ação social da Autoridade para as Condições de Trabalho, assim como tudo o que diga respeito a despedimentos coletivos e falências. Considera-se uma indexação com um grau considerável de especificidade e pertinência, visto ter descritores específicos que abrangem assuntos dominantes. Tratando-se de um

tema atual, percebe-se o cuidado na indexação, o que vai ao encontro da política editorial, visto ter sido publicado numa secção de destaque “Primeiro Plano”.

Ilustração 6 – Documento da imprensa: Exemplo 1



| 2 | JORNAL DE NOTÍCIAS |

| Quinta-feira 8 |
JANERO 2009 |

PRIMEIRO PLANO

AS CONSEQUÊNCIAS DA CRISE ECONÓMICA



Formação profissional no sector automóvel

O processo de regulamentação do Programa Qualificação no Emprego para apoiar o sector automóvel está concluído e que as candidaturas ao mesmo deverão abrir nas próximas semanas.



ACT fiscaliza falências e suspensões de trabalho

Autoridade para as Condições de Trabalho quer assegurar a subsistência dos trabalhadores atingidos

ALEXANDRA FIGUEIRA
afigueira@jn.pt

A Autoridade para as Condições de Trabalho (ACT) vai acompanhar o encerramentos de empresas, despedimentos colectivos e reduções ou suspensões de actividade para assegurar o cumprimento da lei e a protecção dos trabalhadores.

Hoje, os mais de 40 dirigentes locais, regionais e centrais da ACT vão encontrar-se, na habitual reunião de início de ano, mas com um novo ponto na agenda. Antecipando a multiplicação de falências ou redução da actividade de empresas e de despedimentos e suspensão de trabalhadores, o inspetor-geral de trabalho, Paulo Morgado de Carvalho, vai pedir um acompanhamento próximo destes casos “para assegurar que tudo é feito dentro da lei” e que “a subsistência dos trabalhadores – muitos de estratos sociais e económicos não muito altos – é acautelada”, adiantou.

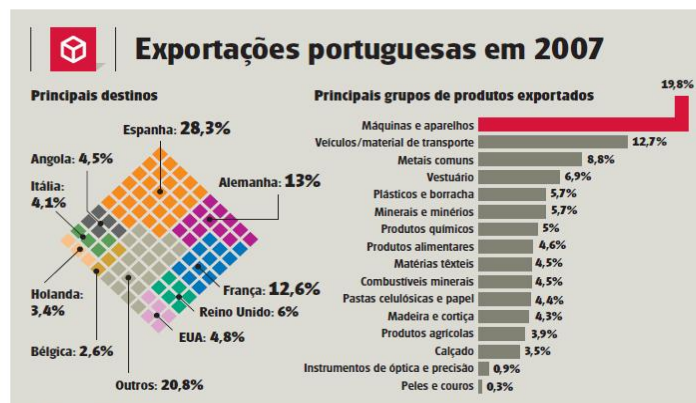
Já ontem a ACT foi ao terreno analisar a suspensão de contrato anunciada pela Tyco Electronics a 536 trabalhadores (ler página 4). Este género de intervenção deverá agora multiplicar-se. “Há em-

Autoridade das Condições de Trabalho quer punir quem use a crise como desculpa para fechar portas

presas a tentar por todos os meios sobreviver, mas, noutros casos, pode haver tentativas de se aproveitarem da conjuntura e nós temos que separar o trigo do joio”.

Legalidade pode dar prisão

A ACT distinguirá as reduções ou suspensões de contrato legais das ilegais, que podem ser punidas com coima ou, até, pena de prisão, lembrou Paulo Morgado de Carvalho. E mesmo entre as feitas nos termos na lei, será necessário acautelar os direitos dos trabalha-



Fuente: NE ICE / Infografia JN



A indústria automóvel em Portugal é uma das que mais tem vindo a reduzir a actividade em face da crise instalada

dores, mas evitando piorar a situação da empresa. “Devemos ter ponderação e bom senso”, disse, para não prejudicar a produção industrial e agravar a crise.

O acompanhamento próximo de falências de empresas ou suspensões de actividade será feito em paralelo à actividade corrente da ACT. “A nossa missão é controlar e fiscalizar, sobretudo no que toca à segurança e ao trabalho ilegal, e não deixaremos em caso algum de o fazer”, assegurou.

Por isso, e no que toca à sinistralidade, por exemplo, a construção civil continuará a receber boa parte das atenções da Autoridade. Mas o acompanhamento dos casos de falências ou redução de actividade, do automóvel (já em graves dificuldades) e outros sectores exportadores e de mão-de-obra intensiva (como o têxtil e o calçado) serão analisados de perto, disse o inspetor-geral.

Mais despedimentos colectivos

O número de pessoas alvo de despedimento colectivo tem aumentado e os últimos dados disponíveis – do primeiro semestre de 2008 – deixam de fora os meses durante os quais a crise agudizou-

Fiscalização às condições de higiene e segurança e ao trabalho ilegal (não declarado) vai continuar

se. Entre Janeiro e Junho, quase 1500 pessoas tinham sido despedidas colectivamente, entre os processos concluídos e os ainda em curso – é mais de metade dos 2625 trabalhadores que perderam o emprego por este via em 2007, de acordo com os dados do Gabinete de Estratégia e Planeamento do Ministério do Trabalho.

Quanto às suspensões ou reduções de actividade (os chamados “lay off”), no primeiro semestre de 2008 atingiram 224 pessoas, face a 604 em todo o ano de 2007. ■



Exemplo 2:

Antetítulo: ENTREVISTA DANUTA
HUBNER, COMISSÁRIA EUROPEIA DA
POLÍTICA REGIONAL

Título: "O grande desafio para Portugal é o
investimento na inovação"

FICHA:

Proprietário: DN

Estado: Arquivado

Data da Notícia: 19/06/2006

Autor: FERNANDO DE SOUSA

CAMPOS DE INDEXAÇÃO

Temas:

**COMISSÃO EUROPEIA
POLÍTICA REGIONAL**

Pessoas:

DANUTA HUBNER

Publicação: DIÁRIO DE NOTÍCIAS
ECONOMIA

Edição: LISBOA

Secção: POLÍTICAS E CONJUNTURA

Página: 6

Análise:

Trata-se de uma entrevista recuperada após uma pesquisa feita com o termo composto "REGIÕES E POLÍTICA REGIONAL". O termo é específico de nível 1, subordinado ao tema específico "Política da Economia" e do termo genérico "Crescimento Económico" do microtesauro "Actividade Económica". Foram aplicados a este conteúdo dois termos temáticos "COMISSÃO EUROPEIA" e "POLÍTICA REGIONAL", um outro descriptor do tesouro Personalidades foi eleito, "DANUTA HUBNER". Este último termo é óbvio, visto se tratar da personalidade entrevistada. Os outros dois foram eleitos a partir do antetítulo, que corresponde à função da entrevistada. Isto sugere que os assuntos tratados durante a entrevista não foram retratados no momento da indexação. Daí concluir que se trata do resultado de uma indexação automática. Também comprova que é necessário ponderação na eleição de termos dos títulos já que poderão não traduzir o verdadeiro assunto do texto, como é o caso. E tal poderá acontecer com entrevistas, cujos títulos são normalmente citações ou, então, em artigos de opinião, já que se tratam de títulos ambíguos e subjetivos, tal como o teor intelectual do respetivo texto. Fica em falta a eleição do descriptor "ENTREVISTA", no campo "Tipo de notícia", visto se tratar de um infra-conceito importante na recuperação de informação desta tipologia textual jornalística. Considera-se uma indexação exaustiva mas pouco pertinente.

Ilustração 7 - Documento da imprensa: Exemplo 2

Diário de Notícias
Economia
| 6 | POLÍTICAS & CONJUNTURA | 2ª FEIRA | 19 JUNHO 2006 |

ENTREVISTA DANUTA HÜBNER, COMISSÁRIA EUROPEIA DA POLÍTICA REGIONAL

“O grande desafio para Portugal é o investimento na inovação”

Danuta Hübner
58 anos
Tem duas filhas
Nasceu a 8 de Abril de 1948 em Nisko, Polónia
Doutorada em Economia pela Universidade de Varsóvia
Entre vários cargos que já ocupou, destaca-se o de conselheira económica do Presidente da Polónia entre 1998 e 2001 e secretária de Estado dos Assuntos Europeus em 2003
É comissária europeia desde 1 de Janeiro de 2004

Uma comissária de trato tranquilo

Danuta Hübner é uma pessoa de trato tranquilo. Quando recebeu o Diário de Notícias, no seu gabinete na sede da Comissão Europeia, já tinha recuperado da noite anterior, onde assistira ao primeiro encontro da sua seleção polaca com a Alemanha. Era, afinal, a primeira prestação da sua Polónia e confessa que viu o jogo com grande ansiedade. Fala em voz relativamente baixa, sempre com a preocupação de ser o mais rigorosa possível.

A comissária da política regional conhece bem a União Europeia. Professora de Economia, acompanhou a Polónia, entre 2001 e 2003, no seu caminho em direção da União Europeia, como secretária de Estado dos Assuntos Europeus.

Nos anos 90 já coordenara as reformas económicas no seu país, que geraram uma das maiores taxas de crescimento da Europa Central. Em 1995, chefiou o processo de negociação para a adesão da Polónia à Organização para a Cooperação e Desenvolvimento Económico (OCDE). A política económica em períodos de transição viria a ser uma das suas especialidades. Também pertenceu à Comissão Económica das Nações Unidas para a Europa.

Desenvolveu a sua formação universitária na Polónia, Espanha, Reino Unido e Estados Unidos.

Danuta Hübner nasceu a 8 de Abril de 1948, na pequena localidade de Nisko, no Sul da Polónia.

(f) Fernando de Sousa
Correspondente em Bruxelas

Portugal continua a ser um dos “bons alunos” da Europa. A comissária responsável pelos assuntos regionais da UE diz que Portugal é dos países que têm “usado os fundos de forma muito eficiente”. A comissária afirma que o alargamento ao Leste é uma oportunidade e considera que Portugal ainda precisa de mais infra-estruturas, daí que a Ota e o TGV sejam vistos com agrado.

Depois da aprovação das Perspectivas Financeiras para o período de 2007-2013, serão em breve apresentadas as principais linhas de orientação para a sua aplicação no que se refere aos fundos estruturais. Quais são as prioridades a seguir nessa aplicação?

Pretendemos centrar-nos em investimentos assentes na Estratégia de Lisboa, apontando para o crescimento, emprego e competitividade.

Na próxima geração dos programas integrados na política de coesão há três prioridades principais: A primeira é a competitividade, inovação e iniciativa empresarial. Tudo o que tenha a ver com o desenvolvimento inovador. Certo tipo de investigação também pode ser financiado pelos fundos estruturais. Em grande parte, estas verbas pretendem estabelecer uma ligação entre a comunidade empresarial e a investigação.

A segunda área prende-se com a sensibilidade das regiões. Ainda há regiões, como em Portugal, que precisam de investimentos adicionais em infra-estruturas físicas. Se não houver boas estradas, o investimento não vem. Portanto, a competitividade destas regiões depende também de infra-estruturas físicas. Porém, há regiões que precisam de mais investimentos em telecomunicações, em banda larga, em tudo o que ajude a desenvolver acessibilidade.

A terceira é o capital humano. Temos de investir em pessoas, não só em formação mas também em educação, de forma a aumentar a empregabilidade das pessoas na Europa.

Bons alunos A comissária responsável pela política regional elogia a aplicação dos fundos feita por Portugal

Temos uma taxa de emprego muito baixa. Quanto mais pobre é o país, mais baixa é essa taxa. Estamos a desperdiçar capital humano na Europa. Por isso, queremos investir em projectos desse tipo.

Temos ainda esta proposta de *Earmarking*, que faz parte da nossa regulamentação, segundo a qual as regiões e os Estados membros são obrigados a investir, em determinadas proporções, em prioridades que estão ligadas à Estratégia de Lisboa.

Mas se a principal prioridade é o investimento em inovação e novas tecnologias, alguns países mais avançados poderão ter melhor capacidade para absorver esses fundos. Isso pode significar que os mais atrasados continuarão com dificuldades na aproximação aos mais avançados...

É por isso que no orçamento europeu há uma parte segundo a qual se pode ter acesso a verbas, com base na excelência. Alguma das universidades, regiões ou Estados membros podem concorrer e ganhar.

Através dos fundos estruturais, o que é uma parte diferente do orçamento, queremos ajudar as regiões europeias mais pobres a terem ca-

pacidade para concorrer em pé de igualdade com as melhores. É por isso que vai poder usar-se os fundos estruturais para aumentar a capacidade em investigação e desenvolvimento.

Qual é a reputação de Portugal na utilização dos fundos? No passado, o ex-presidente da Comissão Europeia, Jacques Delors, considerou Portugal “um bom aluno”. Continuamos a merecer essa classificação?

Estão entre aqueles que têm usado os fundos de forma muito eficiente. Vemos muito valor acrescentado em Portugal. Claro que poderá haver casos particulares, como noutros países, mas não são uma grande preocupação para nós.

O principal desafio para Portugal, hoje em dia é focarem-se mais nos investimentos menos tangíveis, na inovação. Portugal - há como a Espanha mas ao contrário da Irlanda - centrou-se mais, nos últimos vinte anos, nas infra-estruturas físicas, o que podemos compreender. Porém, hoje, Portugal tem de investir fortemente nas capacidades de inovação. Apoiamos vivamente as prioridades que o País nos propôs.

O Portugal é muito diversificado,

“Temos de investir em pessoas, não só em formação mas também em educação. Estamos a desperdiçar capital humano na Europa”

desde a Madeira até Lisboa, dos mais pobres aos mais ricos, ainda o Norte e Alentejo. Têm todos os níveis de desenvolvimento no País. Por isso, têm de investir para se aproximarem no sentido de uma coesão mais equilibrada. É preciso entender que se não investirem em inovação não terão empregos que amanhã sejam seguros, que não vão para a China ou a Índia. É um tipo diferente de competitividade que eriam hoje com o investimento. Com estes investimentos, Portugal ficará em melhores condições de competir na busca de outras fontes de financiamento.

Por último, é preciso investir para inovar. Mas ao mesmo tempo tem de se aplicar restrições para conter o défice público...

É um desafio. De qualquer modo, é preciso investimento público para melhorar a capacidade inovadora. Os fundos têm de ser vistos como um apoio. Mas julgo que Portugal, como muitos outros Estados membros, continua a deparar-se com o desafio da estabilidade e das finanças públicas saudáveis. A longo prazo, a estabilidade macroeconómica ajuda a crescer de uma forma mais sustentável. ■

Exemplo 3:

Ilustração 8 - Documento da imprensa: Exemplo 3

Diário de Notícias

| 7 | OPINIÃO | 2.ª FEIRA | 12. JUNHO 2006 |



Luís Delgado
Jornalista

Uns e outros

É preciso ter muita sorte, diga-se: dois dias antes do primeiro jogo de Portugal, o INE divulga as Contas Nacionais do primeiro trimestre, e não é que o País entrou em milagre? O PIB cresceu, as exportações deram um pulo inesperado, o consumo aumentou ligeiramente e o investimento assim-assim. Melhor seria impossível, tal como o filme.

Embalado pelo clima, pela natureza e pelo futuro, que ninguém sabe o que é, mas também pouco interessa, o País vive eufórico, alheado e preparado para tudo, como escreveu Vasco Pulido Valente, no *Público*: "Se for um triunfo, (os portugueses) esquecem o défice e o IVA e o Governo. Se não for, matam o brasileiro e voltam de férias como de costume. Deprimidos? Qual deprimidos!"

Está dito. Tudo e em poucas palavras.

Ficámos assim? Não, somos assim, desde a fundação...

Não gostamos de nada a prazo, que leve tempo a alcançar, esforço, luta e grande cansaço. Isso é para o outros.

Nós, isolada ou colectivamente, apreciamos o imediatismo, o logo se vê, o anda que se faz tarde.

Scolari, o mestre do optimismo, sempre teve razão quando disse que o Campeonato do Mundo era bastante mais acessível para Portugal e equipas similares.

A confusão de 32 selecções, que vão desde Trindade e Tobago ao Brasil, implica uma disparidade grande entre as capacidades colectivas de cada equipa, embora isso não afaste o único drama deste Mundial: os que chegaram agora, pela primeira vez, vão dar o que não têm, mostrar o que nunca conseguiram, e isso é um perigo permanente.

Quem viu o jogo da Suécia com Tobago? Alguém acreditava que o racionalismo sueco, a sua capacidade em campo, e a facilidade com que jogam não chegaria para dizimar o pequeno país das Caraíbas? Pois, o resultado foi o que se viu.

O mesmo vai acontecer com selecções como a de Angola, Tunísia, Arábia Saudita, Irão, Japão e Coreia do Sul, entre outros.

Teoricamente há um maior desequilíbrio,

entre o grupo dos bons, experientes, maduros e activos, e os que lá estão pela primeira ou segunda vez. O perigo, contudo, é esse. Uns acreditam e invocam os seus louros, e glórias, e os outros, sem nada para reclamar, vão no tudo ou nada, e com grande fé em Deus e na bola.

É a melhor posição, sublinhe-se.

Os melhores, os maiores, os mais falados têm um peso difícil de carregar, uma expectativa elevada, e um problema às costas, que é aguentar a pressão e loucuras nacionais.

Os outros, mais pequenos e sem fantasmas, apresentam-se ansiosos, sim, mas suficientemente tranquilos para marcar pontos. Tudo o que fizerem será sempre mais do que tinham, e melhor do que esperavam.

É com tudo isto que Luiz Felipe Scolari e os nossos 23 jogadores têm de lutar, sendo que para além de Angola ainda teremos de enfrentar o Irão, de que só se sabe que aposta na bomba nuclear, sabe-se lá para quê, e chega à Ale-

manha com a ideia de vencer os cristãos infiéis e maldosos, que querem governar o mundo sem o seu divino e apocalíptico saber. Como seria interessante ter na Alemanha uma banca cheia de *ayatollahs*, mas apenas pacificamente fanáticos pelo futebol.

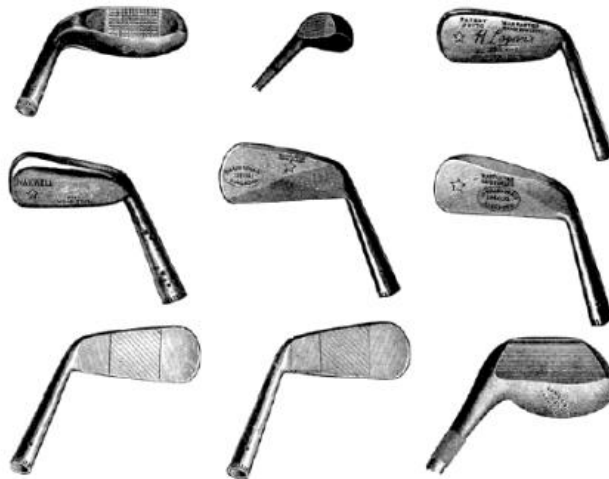
O Mundial a ainda só agora começou, sem nenhuma grande surpresa, mas não terminará sem casos fantásticos, equipas-milagre e heróis mediáticos.

Uma coisa é certa: ninguém está ali a jogar a feijões.

Grandes e pequenos querem a taça que o Pelé empunhou, na cerimónia de abertura, e têm menos de 30 dias para a conquistar. Alguns já mostraram o que valem, como a Alemanha, a Inglaterra e a Argentina.

Mas faltam outros favoritos, como o Brasil, ou aspirantes em grande forma, como a Alemanha, a Croácia, a Ucrânia e alguns mais.

Como vai ser, daqui a 27 dias?!



Os melhores, os maiores, os mais falados têm um peso difícil de carregar, uma expectativa elevada e um problema às costas, que é aguentar a pressão e loucuras nacionais

Maria José Rodrigues

FICHA:

Proprietário: DN
Estado: Arquivado
Data da Notícia: 12/06/2006

Pessoas:

LUIS FILIPE SCOLARI

Instituições:

**INSTITUTO NACIONAL DE
ESTATÍSTICA**

Tipo de Notícia: **OPINIÃO**

CAMPOS DE INDEXAÇÃO

Temas:

**CAMPEONATO DO MUNDO DE
FUTEBOL**

CONTABILIDADE NACIONAL

Dados de Publicação

Publicação: DIÁRIO DE NOTÍCIAS

Edição: LISBOA

Secção: OPINIÃO

Página: 7

Análise:

Trata-se de um artigo de opinião recuperado após uma pesquisa feita com o termo composto (expressão adjetival) “CONTABILIDADE NACIONAL”. O termo é genérico do termo de topo ACTIVIDADE ECONÓMICA, presente no tesouro temático. Foram aplicados dois termos temáticos “CONTABILIDADE NACIONAL” e “CAMPEONATO DO MUNDO DE FUTEBOL”, sendo de áreas tão distintas mas que, neste caso, são dois assuntos que se cruzam, já que o autor/jornalista aborda paralelamente a situação das contas nacionais no início do referido campeonato do mundo. É de realçar que “Contas nacionais” é um não-descritor e sinónimo do primeiro termo indexado. Um outro descritor do tesouro Personalidades foi eleito “LUÍS FILIPE SCOLARI” assim como um do tesouro Intituições “INSTITUTO NACIONAL DE ESTATÍSTICA”. Estes termos não assumem assunto no artigo de opinião, apenas servem de argumentos para a sustentação da mesma, os quais poderiam ser dispensáveis. Neste caso, apenas o conteúdo do artigo poderia ser alvo de análise já que o título é bastante ambíguo, dada a natureza do documento. Parece notório que um texto de opinião terá, necessariamente, uma abordagem mais sensível e atenta no momento da indexação, pois a linguagem é menos objetiva e factual. Realça-se a eleição do descritor infraconceito “Opinião” no campo “Tipo de notícia”. Considera-se uma indexação exaustiva e com pouca pertinência, provavelmente resultante de uma indexação automática.

Exemplo 4:

Ilustração 9 - Documento da imprensa: Exemplo 4

2 | JORNAL DE NOTÍCIAS

Sábado 22
AGOSTO 2009

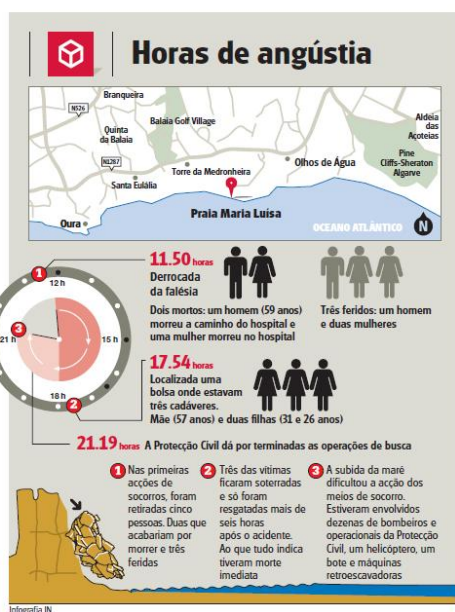
PRIMEIRO PLANO

ACIDENTE NAS ARRIBAS ALGARVIAS



"A sombra acabou em tragédia"

Manuel Castro estava a sair da água quando ouviu um estrondo. "As pessoas entraram todas em pânico. Ouviu-se alguém a gritar". E o metalúrgico de Guimaraes, de 46 anos, lançou-se para a zona sinistrada, como muitos outros populares, na ansia desesperada de



Subida da maré dificultou as operações de resgate dos corpos que só terminaram cerca das 21.20 horas

Tragédia mata família no último dia de férias

Derrocada de arriba na praia Maria Luísa, em Albufeira, no Algarve, fez cinco vítimas mortais

PEDRO IVO CARVALHO*
pic@jn.pt

O desmoronamento de uma arriba na praia Maria Luísa, em Albufeira, Algarve, matou cinco pessoas. Quatro eram da mesma família e viviam no Porto. A quinta era uma mulher de 37 anos, de Coimbra. Deixa uma filha de 11 anos.

As causas da tragédia, que provocou ainda três feridos - dois ligeiros (uma mulher de 30 anos e uma jovem de 16) e um grave (um homem de 24 anos) -, estão ainda por apurar, mas, segundo informações avançadas pelas autoridades marítimas pode ter havido uma conjugação de dois factores: por um lado, o sismo que na terça-feira se fez sentir na região do Algarve (4.2 na escala de Richter);

por outro, as marés vivas que estão a atingir a costa algarvia.

No último dia de férias, a tragédia abateu-se sobre a família Fonseca, residente na freguesia de Ramalde, Porto: quatro dos cinco mortos são marido, mulher e duas filhas. António José Baptista Mota da Fonseca, Anabela Fonseca, Rita Fonseca e Mariana Fonseca tinham, respectivamente, 59 anos, 57, 31 e 26. A família encontrava-se há cerca de duas semanas de férias na zona de Albufeira, na casa de José Dantas, conhecido médico do Porto, contou, ao JN, Manuel Maio, presidente da Junta de Freguesia de Ramalde e amigo da família e do médico.

A quinta vítima mortal é uma mulher de 37 anos, residente em Coimbra. Foi resgatada das pedras

populares ainda com vida e viria a falecer no Hospital de Faro. António Fonseca ainda foi reanimado no local, mas viria a morrer na ambulância, vítima de paragem cardiopulmonar.

O ferido grave, Vitor Sousa, de 24 anos, era namorado de uma das filhas de António Fonseca, ex-director desportivo do Boavista. O jovem foi sujeito a uma intervenção cirúrgica a uma perna no Hospital de Faro e não tinha sido, até cerca das 23 horas, informado da morte da namorada e da respectiva família. "Ele já perguntou por eles e achou estranho ninguém ir visitá-lo ao hospital", confidenciou, ao JN, Manuel Maio.

Policia atendeu telemóvel

A notícia da tragédia chegou a Ra-

Uma família simpática

Na Rua da Venezuela, no Porto, onde morava a família Mota da Fonseca, o 4.º esquerdo do número 276 tem as persianas abertas como se lá estivesse gente. Puro engano. "Eram pessoas simpáticas, educadas e muito agradáveis", explicou, ao JN, Francisco Caldas, administrador do condomínio, visivelmente transtornado. A família sempre viveu naquela rua, num dos prédios ao lado, mas há três anos mudou de casa. "São as únicas referências que tenho", resumiu o vizinho. **N.A.L.**

malde por uma funcionária de uma instituição onde era educadora de infância Anabela Fonseca. Logo Manuel Maio, conhecido do facto de a família estar de férias no Algarve, naquela zona, temeu o pior.

Ligou para os telemóveis das quatro vítimas. Três não davam sinal. E apenas o telefone de António Fonseca tocou. Aliás, terá mesmo sido atendido por um polícia quando a vítima estava a ser transportada para o hospital.

"Foi aí que percebemos a dimensão da tragédia", conta Manuel Maio, que trabalhou de perto com duas das vítimas: a mulher, educadora de infância na Junta, e uma das filhas, Rita, psicóloga na mesma autarquia. A irmã, Mariana, era jurista.



FICHA:

Proprietário: JN

Estado: Arquivado

Data da Notícia: 22/08/2009

Antetítulo: ACIDENTE NAS ARRIBAS
ALGARVIAS

Título: Tragédia mata família no último dia de
férias

Autor: PEDRO IVO CARVALHO

GINA PEREIRA

CAMPOS DE INDEXAÇÃO:

Temas: **DERROCADAS**

SOTERRADOS

Localidades: ALBUFEIRA

Publicação: JORNAL DE NOTÍCIAS

Edição: PORTO

Secção: PRIMEIRO PLANO

Página: 2

Análise:

Este documento foi recuperado a partir de uma pesquisa com o termo “Condições de Trabalho”. Tratando-se de um termo genérico, todos os documentos indexados com os respetivos descritores hierarquizados como específicos daquela árvore temática são recuperados. Espera-se que desta pesquisa resulte documentos subordinados a tudo o que se relacione com as condições laborais. Na verdade, a esmagadora maioria dos 830 documentos recuperados vão ao encontro dessa temática. Porém, foi perceptível a existência de alguns que não tinham esse enquadramento. É o caso deste documento, que foi indexado com os termos “DERROCADA” e “SOTERRADOS”. Denote-se que este último descritor é um termo específico de nível 2, subordinado ao TE1 “Acidente de Trabalho” e ao TE “Segurança no trabalho”. Percebe-se, então, que o termo “Soterrados” é aqui aplicado noutra contexto, ou seja, numa notícia sobre acidente em férias. Visto que este sistema integrado não permite relacionar termos nem repeti-los noutra árvore, o mais importante no momento da indexação será sempre recuperar tudo o que se relaciona com o termo “Soterrados”, independentemente de estar, ou não, devidamente contextualizado com o enquadramento do tesauro. Impõe-se, assim, aplicar um determinado termo para depois recuperá-lo, sendo um mal menor o facto de provocar ruído na recuperação da informação. Apesar deste exemplo constituir um caso residual, com ele reforça-se a relevância da proximidade da representação do assunto ao jargão jornalístico e, conseqüentemente, às necessidades do público principal deste arquivo, que são os jornalistas. Trata-se, portanto, de uma indexação específica e pertinente, tendo em conta o grande objetivo da recuperação da informação.

Exemplo 5:

Comunistas PCP diz que retoma só se vê «à lupa»

> RUI CUNHA

Aveiro

Carlos Carvalhas lembrou ontem em Aveiro que «todos os índices publicados pelo Instituto Nacional de Estatística até agora mostram que não tem havido retoma». «À lupa é possível que se veja, porque ela é irrisória», ironizou o secretário-geral do PCP após um encontro com a reitora da Universidade de Aveiro.

«Nenhum país pode estar sempre a decrescer e portanto há alturas em que basta subir um bocadinho para já haver retoma em relação ao sítio em que se está», afirmou o comunista, advertindo, porém, que Portugal irá continuar a afastar-se da média europeia «até pelo menos 2006», apesar de «o primeiro-ministro ter prometido que com o seu Governo o país crescerá sempre dois pontos acima da média europeia». Carlos Carvalhas defende uma «estratégia de desenvolvimento totalmente diferente», uma vez que a «política actual de contenção cega, com quebra do investimento público, leva à recessão».

O líder do PCP disse ainda que o «descontentamento geral» que se sente no país poderá ter consequências na adesão dos cidadãos às eleições para o Parlamento Europeu.

Ilustração 10 - Documento da imprensa: Exemplo 5

FICHA:

Proprietário: DN
Estado: Arquivado
Data da Notícia: 09/03/2004
Título: Comunistas PCP diz que retoma só se vê «à lupa»

Autor: RUI CUNHA_AVEIRO

CAMPOS DE INDEXAÇÃO

Temas:

RETOMA ECONÓMICA

ESTATÍSTICAS NACIONAIS

PARTIDO COMUNISTA PORTUGUÊS

Localidades: **AVEIRO**

Dados de Publicação

Publicação: DIÁRIO DE NOTÍCIAS

Edição: PORTO

Secção: NACIONAL

Página: 5



Análise:

Foi feita uma pesquisa com o termo composto (expressão adjetival) “ESTATÍSTICAS NACIONAIS”. Trata-se de um termo específico de nível 1, subordinado ao termo específico “Estatística “ e ao termo genérico “Análise Económica”. Esta estrutura implica a indexação de documentos que expliquem uma abordagem analítica no domínio económico, nomeadamente tudo o que diga respeito a estudos estatísticos portugueses. Este termo também se encontra no tesouro “TIPO”. Tal acontece para que todos os documentos que apresentem dados estatísticos sejam recuperados, não surgindo como assunto desses mesmos documentos. No entanto, também poderá surgir como tema principal, embora aconteça com uma menor frequência. Foram aplicados três termos como tema “RETOMA ECONÓMICA”; “PARTIDO COMUNISTA PORTUGUÊS” e “ESTATÍSTICAS NACIONAIS” e um geográfico “AVEIRO”. O terceiro termo surge como *infraconceito* visto se tratar de uma notícia relativa à publicação de dados estatísticos. O segundo refere-se a quem protagonizou a análise dos resultados estatísticos. O primeiro é o assunto crucial da notícia. O descritor geográfico especifica o local onde ocorreu a base factual. Desta forma, a indexação realizada responderá às questões *o quê?*; *quem?* e *onde?*, tendo sido utilizados dois dos tesouros existentes: temas e geográfico. Esta notícia pode ser recuperada numa pesquisa sobre estatísticas nacionais e também sobre a retoma económica. Para além disso, é pesquisável para quem pretender recuperar documentos relativos à ação política do Partido Comunista Português. Por fim, agrupa-se em todos os conteúdos recuperáveis pelo distrito/cidade de Aveiro. Considera-se uma indexação exaustiva e com pertinência.

Exemplo 6 :

2 | JORNAL DE NOTÍCIAS

Ilustração 11 - Documento da imprensa: Exemplo 6

Terça-feira 3
FEVEREIRO 2009

PRIMEIRO PLANO

SONDAGEM UNIVERSIDADE
CATÓLICA/JN/RTP/ANTENA1



Os admiradores do primeiro-ministro

Não são muitos, mas não deixa de ser significativo que haja quem tenha mudado a sua opinião sobre Sócrates para melhor (8%). Incluindo entre os que simpatizam com a oposição (4%).



Explicações de Sócrates não foram suficientes

Sondagem revela que a maioria dos portugueses entende que o primeiro-ministro deixou vários aspectos por esclarecer no que diz respeito ao seu envolvimento no processo Freeport



Primeiro-ministro fala de uma campanha negra. E muitos dos inquiridos (38%) concordam

RAFAEL BARBOSA
rafael@jn.pt

O processo Freeport está a abrir brechas na credibilidade de José Sócrates. A esmagadora maioria dos inquiridos na sondagem (61%) acham que o primeiro-ministro deixou coisas por esclarecer. E uma percentagem significativa (43%) não acredita que não tenha havido favorecimento.

José Sócrates já convocou a Comunicação Social por duas vezes, para falar sobre o processo Freeport. A primeira, há pouco mais de uma semana, na Alfândega do Porto. A segunda, no Palácio de S. Bento, residência oficial do primeiro-ministro, na quinta-feira passada. No entanto, e a julgar pela sondagem telefónica que a Universidade Católica realizou durante o último fim-de-semana, José Sócrates não convenceu os portugueses. Entre os 84% de inquiridos que ouviram falar do caso Freeport, a esmagadora maioria (61%) entende que o primeiro-ministro deixou coisas por esclarecer. Apenas uma pequena franja (18%) se sente completamente esclarecida.

Pior ainda, são também em maior número os portugueses que entendem que houve alguma espécie de favorecimento no processo de aprovação do Freeport. Quando confrontados com a afirmação de José Sócrates, segundo o qual não houve favorecimento, são 43% os que dizem não acreditar no primeiro-ministro. Quase o dobro dos que acreditam (23%). Sendo que um terço dos inquiridos (32%) não é capaz de opinar sobre matéria tão sensível.

Conforto socialista

Quando se cruzam estas questões com a identificação partidária dos inquiridos, os resultados são diferentes. Entre quem não tem simpatia por qualquer partido, como entre

os que a têm por partidos de oposição, a desconfiança face a José Sócrates aumenta. Mas mesmo entre os simpatizantes socialistas se nota que a imagem do também secretário-geral do PS está fragilizada. Se é verdade que 45% dos inquiridos se sentem esclarecidos, são quase tantos (37%) os que não estão. Os que acreditam que não houve favorecimento são 48%, mas ainda há 21% que respondem em sentido contrário. Resumindo, apenas metade dos simpatizantes socialistas parece manter confiança absoluta no líder do PS.

Actuação das autoridades judiciais não agrada

O principal personagem do caso Freeport pode ser José Sócrates, mas a justiça também está, por estes dias, sob escrutínio. Sendo que a avaliação que os portugueses fazem não é positiva. São mais (42%) os que acham que as autoridades judiciais portuguesas têm agido mal ou muito mal, do que aqueles que acreditam na capacidade dos nossos polícias e magistrados (34%). E, nesta matéria, a simpatia partidária tem muito pouco peso. Se a actuação das autoridades judiciais não tem sido a melhor, segundo os inquiridos, também a crença no resultado final teria de ser negativa. São 48% os que não têm confiança que a investigação esclareça rapidamente todas as dúvidas. Ainda assim, 42% ainda acredita na máquina judicial. E, aqui sim, os simpatizantes do PS são os que revelam maior optimismo (61%).

Tendo em conta as respostas anteriores, facilmente se concluiria que a opinião sobre José Sócrates não seria a melhor por estes dias. Embora 50% não tenha registado mudanças (a sondagem não esclarece se a opinião era boa ou má), são muitos mais os que apontam uma mudança de opinião para pior (31%) dos que os que escolhem para melhor (8%).

Campanha negra

No que José Sócrates parece estar a ser relativamente eficaz é na retórica sobre a "campanha negra" de que diz estar a ser alvo. A maior fatia de inquiridos (38%) está de acordo, embora se registre um empate técnico (36%) com quem não subcreve tal argumento. E, aqui sim, funciona em pleno a solidariedade partidária, uma vez que são 72% os simpatizantes socialistas que também alinham na tese da cabala.

Os dotes comunicacionais de José Sócrates ficam claramente confirmados nesta sondagem. Porque, independentemente do que pensam sobre o processo Freeport e sobre o envolvimento directo do pri-

Sócrates disse que não houve qualquer espécie de favorecimento. A maior parte (43%) não acredita

meiro-ministro, os inquiridos não penalizam a sua actuação neste assunto. Se excluirmos os que não sabem ou não responderam, a amostra divide-se rigorosamente ao meio entre os que acham que Sócrates tem agido bem ou muito bem e os que acham que tem agido mal ou muito mal. ■



COMENTE ESTA SONDAÇÃO
www.jn.pt/



FICHA:

Proprietário: JN

Estado: Arquivado

CAMPOS DE INDEXAÇÃO

Temas:

SONDAGEM

CASO FREEPORT

Pessoas:

JOSÉ SÓCRATES

Dados de Publicação

Publicação: JORNAL DE NOTÍCIAS

Edição: PORTO

Secção: PRIMEIRO PLANO

NACIONAL

Página: 2

Análise:

Foram aplicados dois termos como tema “SONDAGEM” e “CASO FREEPORT” e um de personalidades “JOSÉ SÓCRATES”. Este termo responde à questão “*quem?*” e é o cerne da sondagem realizada a propósito do Caso Freeport, que é o segundo descritor temático e que corresponde à questão “*o quê?*”, que bem denota a presença da linguagem natural no termo controlado. Todos eles são eleitos a partir da sua presença nos títulos e na frequência do texto, o que permeabiliza a ajuda da indexação automática. O primeiro termo surge como *infraconceito* visto se tratar de uma notícia relativa à análise de uma sondagem, embora aplicado no campo “Temas”. Esta notícia pode ser recuperada numa pesquisa sobre sondagens e também sobre o Caso Freeport. Para além disso, é pesquisável para quem pretender recuperar documentos relativos à ação política de José Sócrates. Considera-se uma indexação com um grau considerável de exaustividade e pertinência, dada à atualidade e destaque do assunto.

Exemplo 7:

Ilustração 12 - Documento da imprensa: Exemplo 7

38 | dn | QUARTA-FEIRA 7 JULHO 2004 | NEGÓCIOS

Interbancário 1 dia 2.40 2.15 1.90 1.65 1.40 05/06 06/07 2,3994	Euribor A SES MESES 2.240 2.215 2.190 2.165 2.140 05/06 06/07 2,166	Euro/Dólar 1.235 1.225 1.215 1.205 1.195 05/06 06/07 1,2295	Petróleo 38,00 36,75 35,50 34,25 33,00 05/06 06/07 37,18
---------------------------------------------------------------------------------------------	-------------------------------------------------------------------------------------------------	-----------------------------------------------------------------------------------------	--------------------------------------------------------------------------------------

INQUÉRITO

Portugueses estão a poupar menos desde início do ano

Arquivo DN/2003-Paulo Sprangor



Homens entre os 25 e 34 anos, das classes alta e média alta, são os que mais pouparam

■ CÁTIA ALMEIDA
Os portugueses estão a poupar menos este ano do que em 2003. Um fenómeno que pode estar relacionado com a tão esperada retoma económica, revela um estudo realizado pela Marktest.

Segundo os resultados da primeira vaga de 2004 da pesquisa, a Basef Banca, o número de portugueses com conta bancária que no último ano destinaram alguma percentagem dos seus rendimentos à poupança situa-se nos 26,5%. No mesmo período do ano transacto este valor atingia os 32,5%. Ou seja, um decréscimo de 18,5%.

Contudo, há mais portugueses que estão a investir significativa-

mente naquele objectivo. Por outras palavras, no inquérito realizado este ano há mais indivíduos que afirmam destinar maiores percentagens à poupança.

No global, a maioria das pessoas que «armazenam» parte do rendimento diz que está a poupar menos. Na primeira vaga de 2003, 54,6% fizeram esta afirmação, enquanto em 2004 este valor registou uma subida para os 58,5%.

De acordo com a Marktest, os mais poupados são os homens (com 32,1% a afirmar que destina algum dinheiro para esse objectivo), entre os 25 e 34 anos (37,4% poupam), pertencentes às classes alta e média alta (40,7%) e ocupando funções de quadros médios superiores (49,3%).

Se as duas últimas características sócio-económicas não representam qualquer surpresa, o mesmo não se pode dizer em relação ao sexo e à idade.

A amostra do estudo foi constituída por 15000 entrevistas a indivíduos com mais de 15 anos e detentores de conta bancária. Todos os concelhos do País estão representados, sendo a recolha dos dados dividida em três vagas.

O Basef Banca é um estudo regular lançado em 1989 e que tem por objectivo apoiar o marketing dos bancos, fornecendo informação sobre os comportamentos, atitudes e opiniões dos consumidores efectivos e potenciais dos serviços financeiros destinados aos particulares.

QUEDA. Percentagem dos indivíduos que pouparam passou de 32% para 26%



FICHA:

Proprietário: DN

Estado: Arquivado

Autor: CÁTIA ALMEIDA

CAMPOS DE INDEXAÇÃO

Temas:

POUPANÇA

Dados de Publicação

Publicação: DIÁRIO DE NOTÍCIAS

Edição: LISBOA, PORTO

Secção: NEGOCIOS

Página: 38

Análise:

Foi aplicado um único termo como tema “POUPANÇA”. Trata-se de um termo específico de nível 1, subordinado ao termo específico “Rendimento” e ao termo genérico “Contabilidade Nacional”. Este termo responde à questão “o quê?” e é o cerne do inquérito realizado a propósito dos hábitos de poupança. Foi eleito a partir da sua presença nos títulos e na frequência do texto. Esta notícia pode ser recuperada numa pesquisa sobre poupança. No entanto, não foi aplicado o descritor infraconceito “inquérito”, para que estivesse recuperável para quem quisesse pesquisar tudo o que se relacionasse com inquéritos realizados. Considera-se uma indexação com grande nível de especificidade e pertinência, relacionando-se com a secção “Negócios”, onde está publicado o documento.

Exemplo 8

Protestos de pouco serviram



Ilustração 13 - Documento da imprensa: Exemplo 8

Cinco Urgências nocturnas encerram até Junho

Viana do Castelo. Fecho dos SAP está dependente da abertura de novos serviços

Durante o primeiro semestre deste ano, cinco Serviços de Atendimento Permanente (SAP) deverão encerrar, no período nocturno, no distrito de Viana do Castelo, processo que apenas está dependente da conclusão da instalação dos dois novos Serviços de Urgência Básica (SUB). A garantia foi avançada pelo coordenador da Sub-Região de Saúde de Viana do Castelo.

Os SUB serão instalados nos Centros de Saúde de Monção e Ponte de Lima. "Estão em fase de conclusão do projecto e as obras deverão demorar cerca de três meses a concluir. Só depois é que poderemos avançar com os

encerramentos nocturnos dos SAP, talvez ainda no primeiro semestre de 2008", admitiu João Carneiro.

Depois de meses de protestos de rua, a irreversibilidade do ministro da Saúde venceu os autarcas contestatários do encerramento dos SAP. Inicialmente, o prazo para o encerramento era Março de 2008, data que, ao que tudo indica, não será cumprida, "porque não encerrará nenhum SAP enquanto não estiverem implementados todos os meios e equipamentos que respondam com eficácia a situações urgentes e de emergência dos doentes", garantiu João Carneiro.

Em Valença, milhares de populares chegaram a cortar a ponte que liga a Espanha, protestando contra o fim das "urgências" nocturnas e a decisão de deslocar o SUB para Monção, a 20 quilómetros. O SAP fechará às 00.00, contando 24 horas com uma ambulância de Suporte Intermédio de Vida (SIV). Já em Arcos de Valdevez, mais de cinco mil pessoas contestaram na rua a decisão de encerramento do SAP, a partir da 22.00, mas sem resultados. O concelho ficará igualmente servido com uma ambulância de SIV.

A Administração Regional de Saúde do Norte homologou protocolos também com as câmaras de Melgaço, Paredes de Coura e Caminha, para o respectivo encerramento nocturno dos SAP.

No âmbito da reestruturação dos cuidados primários de saúde, o distrito vai ainda contar com uma terceira ambulância SIV em Viana do Castelo (durante o dia), seguindo para Melgaço, no extremo do concelho, no período nocturno, para além do apoio de um helicóptero do INEM. ■ - P.J.

PROTESTOS CONTRA URGÊNCIAS

2008 começou com manifestações nas regiões norte e centro. Hoje continuam

O primeiro dia do ano ficou marcado por vigílias de utentes, em protesto contra o encerramento dos serviços de atendimento permanente - SAP - de Vouzela e S. Pedro do Sul e contra as Urgências do Hospital de Anadia. O protesto consistirá numa concentração nas Urgências para serem atendidos, mas sem prejudicar aqueles que lá se possam deslocar com necessidades reais. Os protestos prosseguem hoje.

FICHA 1

Proprietário: DN

Estado: Arquivado

Data da Notícia: 02/01/2008

Título: Cinco Urgências nocturnas encerram até Junho

Pós-Título: Viana do Castelo. Fecho dos SAP está dependente da abertura de novos serviços

Legenda: Protestos de pouco serviram

CAMPOS DE INDEXAÇÃO

Temas: ENCERRAMENTO DE URGÊNCIA HOSPITALAR

ENCERRAMENTO DE SAP

Localidades: VIANA DO CASTELO

MONÇÃO

PONTE LIMA

Observações: CRONOLOGIA DO MÊS

Publicação: DIÁRIO DE NOTÍCIAS

Edição: LISBOA

Secção: PORTUGAL

Página: 1

FICHA 2:

Proprietário: JN

Estado: Arquivado

Data da Notícia: 18/08/2007

Antetítulo: TRADIÇÃO POPULAR

Título: Ouro brilhou pelas ruas

Autor: ANA PEIXOTO FERNANDES

CAMPOS DE INDEXAÇÃO

Localidades: VIANA DO CASTELO

Publicação: JORNAL DE NOTÍCIAS

Edição: PORTO

Secção: PRIMEIRO PLANO

Página: 2

TRADIÇÃO POPULAR



Ouro brilha pelas ruas

- ▶ Cerca de 400 mulheres no desfile que marca o arranque da romaria da Senhora da Agonia
- ▶ Mordomas de Viana do Castelo com trajes de lavadeiras acompanhadas pelos zabumbas

Ana Peixoto Fernandes
 Alfredo Cunha

Foram perto de quatrocentas as raparigas de todas as idades, profissões e estratos sociais que envergaram, ontem, os trajes regionais de Viana do Castelo, desfilando pelas ruas da cidade, espalhando cor e brilho – do ouro e da sua própria beleza. Foi o Desfile da Mordomia que, todos os anos, inaugura oficialmente a Romaria de Nossa Senhora da Agonia. Como sempre, o cortejo atravessa a cidade apinhada de gente deslumbrada com o imenso colorido estampado na rua pelas centenas de trajes, para apresentar os cumprimentos às autoridades: Governo Civil, Câmara e Diocese.

“Este é um dos actos mais simbólicos que se conhece. Vemos muitas centenas de raparigas envergando trajes únicos, que têm um grande significado para todos nós”, refere o orientador do cortejo, Alberto Rego, homem ligado ao folclore e às tradições da região, que acrescenta: “Estão aqui raparigas das várias classes sociais, algumas muito bem na vida, porque têm os seus cursos superiores, são professoras, doutoras, mas também raparigas mais modestas, que trabalham no campo, em casa ou que ainda procuram o primeiro emprego.

A movimentação das mordomas começa bem cedo. Logo pela manhã e é vê-las chegar sorridentes, aos magotes, ao largo do edifício do Governo Civil. Os vermelhos, verdes e azuis predominam na ala das “moçoilas”, que envergam fatos de lavadeira. O vermelho é a cor

Vozes estrangeiras



Tinie Van Den Wildenberg
65 anos, holandesa

Há quatro semanas em Viana do Castelo, Tinie tem espelhado na cara o deslumbramento com uma multidão de turistas que nesta altura do ano vivem as festas da cidade. A quantidade de ouro que vê ao peito das raparigas trajadas com fatos regionais espanta-a. “Não compreendo como é que as jovens têm tanto ouro”, exclama. O resto é alegria: “O tempo está bom, o que mais queremos?”.



Stephanie Promberger
23 anos, austríaca

A jovem austríaca é uma repetente nas festas de Viana. “Tentamos vir todos os anos e adoro”, diz. E, afinal, o que tanto atrai nestas festividades? “O convívio, o orgulho das pessoas de participar e a música, completamente diferente da tradição austríaca”, responde. A emoção de viver estes dias é tanta que, afirma: “Até acordo cedo para vir”.

principal, sejam os trajes típicos à vianesa, de Areosa, Afife ou Carreço. Depois também os há na versão azul, e os verdes tradicionais da freguesia de Geraz do Lima. Os pretos são, contudo, as atracções do momento. As vestes de Mordoma honram a rapariga que as trajam. Por isso, desfilam como rainhas, ostentando lenços de todas as cores e pesadas peças de ouro ao peito.

A quantidade de máquinas fotográficas por metro quadrado revelou bem o motivo da enchente de ontem na cidade. “Acho que este ano as Mordomas estão mais bonitas. Talvez porque se dispuseram de forma mais organizada. Estavam muito airozas, abertas, organizaram-se muito melhor. Tive muito prazer em colaborar com estas raparigas, porque correu mesmo muito bem”, afirmou Alberto Rego, como que a justificar o brilhantismo do cortejo e as longas filas de gente acumuladas ao longo dos passeios e praças da zona urbana.

No entusiasmo provocado pela passagem da mordomia, só os homens dos bombos conseguiram destronar as protagonistas. Em plena Avenida dos Com-

batentes, a principal artéria da cidade, milhares de curiosos aplaudiram as exibições dos grupos de zabumbas, como que a desafiar-las a tornar ainda mais ensurdecedor o ribombar, tão típico das festas de Viana do Castelo.

A romaria prossegue até à próxima segunda-feira e os trajes regionais prometem voltar a brilhar já na noite de hoje.

“Amanhã [hoje] vamos à Festa do Traje. Lá explicam tudo sobre os fatos”, comenta uma jovem espectadora do Desfile da Mordomia para o namorado. A festa tem lugar às 22 horas no Castelo de Santiago da Barra. <

Vermelho é a cor predominante. Mas os trajes pretos são a atracção do momento

Análise:

Estes dois documentos, que constituem de uma forma conjunta o exemplo 8, foram recuperados com a pesquisa a partir do termo “VIANA DO CASTELO”, descritor oriundo do tesouro geográfico. Daqui espera-se encontrar documentos que representem factos ocorridos nesta localidade, o que de facto aconteceu. Porém, parece pertinente abordar uma constante na indexação que incide no critério de interesse, tendo em conta a sua dimensão e a política editorial. Sempre que determinado conteúdo seja considerado de interesse nacional, encontramos uma indexação mais exaustiva, como é exemplo o documento 1, com os termos “VIANA DO CASTELO”; “MONÇÃO”; “PONTE DE LIMA”; “ENCERRAMENTO DE URGÊNCIA HOSPITALAR” e “ENCERRAMENTO DE SAP”. Tratando-se de um assunto de amplitude nacional, como foi o encerramento das urgências hospitalares, houve a necessidade de alargar a indexação. Já no exemplo 2, denota-se uma grande especificidade, sendo eleito apenas o descritor geográfico, pois não se considera de interesse nacional as Festas da Senhora da Agonia. Desta forma, quanto ao primeiro documento, trata-se de uma indexação exaustiva e pertinente. Já o segundo, prima pela especificidade, não deixando, porém, de ser pertinente. Todavia, poder-se-ia acrescentar um dos descritores temáticos “Festas e Romarias”; “Festas populares” ou “Festas religiosas” para que fosse possível recuperar por esta área temática. Assim, este documento ficará ausente desse assunto, provocando um nível de silêncio na recuperação. Neste aspeto, daqui se depreende que a política editorial impõe-se na eleição dos descritores, interferindo na política de indexação.

3.1.1- Análise geral de 50 documentos

Para além destes casos, 50 outros documentos foram igualmente analisados (ANEXO 3), donde foi possível delinear as variáveis base da política de indexação dos conteúdos jornalísticos da imprensa do Grupo Controlinveste. Denote-se que não é pretensão fazer destes exemplos uma amostra, mas apenas uma demonstração aleatória, em que se perceba as tendências da indexação deste tipo de conteúdos. Contudo, houve o cuidado de selecionar documentos oriundos de diferentes secções dos jornais “Jornal de Notícias” e “Diário de Notícias” de forma a abranger diferentes áreas temáticas e tipologias, aos quais foi feita uma análise quantitativa e qualitativa da indexação efetuada, tendo em conta as seguintes variáveis e objetivos:

Variáveis	objetivo
A Nº de descritores por documento	- Verificar o nível de especificidade ou exaustividade, considerando esta última a partir de 3 descritores por documento.
B Utilização de Tesouros por documento	- Perceber qual/quais o/s tesouro/s mais utilizado/s na indexação e as áreas temáticas mais pertinentes.
C Recurso ao infraconceito	- Compreender a necessidade da utilização do infraconceito como denominador da tipologia textual.
D Harmonia do assunto com a secção onde o documento está inserido	- Verificar se a indexação está ajustada à estrutura intelectual do jornal.
E Relação dos descritores com a linguagem natural do documento	- Aprender a ligação dos descritores (linguagem controlada) com a linguagem natural dos documentos (títulos, legendas, <i>lead</i>).

Tabela 12 – Variáveis e objetivos da análise dos documentos da imprensa

A-Número de descritores por documento

Considerando que a indexação tem um nível de especificidade quando são aplicados até dois descritores por documento, sendo três ou mais indicação de uma indexação exaustiva, do total dos 50 documentos analisados, 27 foram indexados com um máximo de dois descritores, enquanto que 23 têm 3 ou mais descritores. Por esse motivo, apreende-se que há uma tendência para que a indexação seja mais específica do que exaustiva, embora os valores sejam muito próximos. Daqui perfaz que 54% dos conteúdos têm uma indexação com nível de especificidade e 46% com exaustividade.

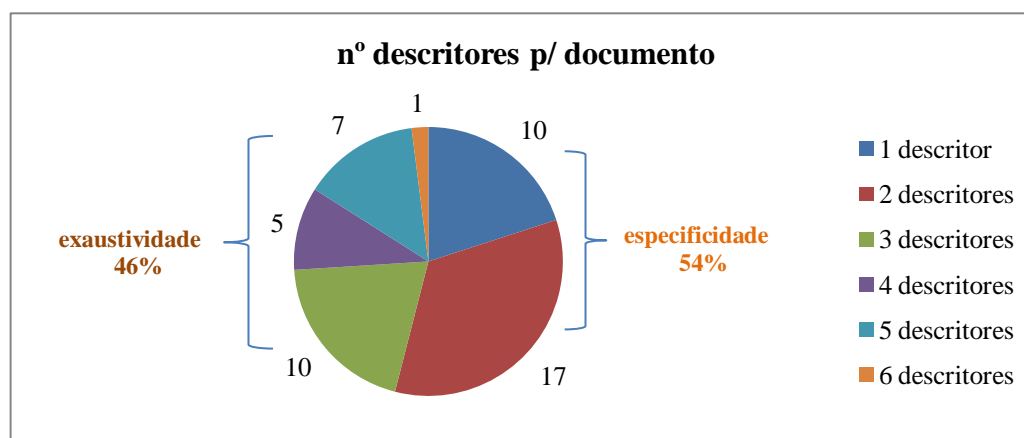


Gráfico 2 – Nº de descritores por documento: exaustividade / especificidade

B-Utilização de Tesouros por documento

Dos 50 documentos analisados, 22 documentos foram indexados com descritores de um tesouro (44%), 17 com descritores de dois (34%) e 11 com descritores de três tesouros (22%). Dos 6 tesouros utilizados (Temas; Instituições; Empresas; Pessoas e Geográfico), 46 documentos foram indexados com descritores do tesouro “Temas”, 15 com o “Geográfico”; 11 com o “Tipo”; 6 com descritores do tesouro “Empresas” e 5 com “Instituições, o que se percebe a predominância da indexação pelos seus temas, podendo esta ser auxiliada com a localização geográfica e/ou com a tipologia documental. Claro está que a utilização de nomes próprios, relativos a pessoas singulares ou coletivas, tem pertinência neste tipo de conteúdo, visto serem protagonistas de notícias publicadas, daí a sua alargada utilização.

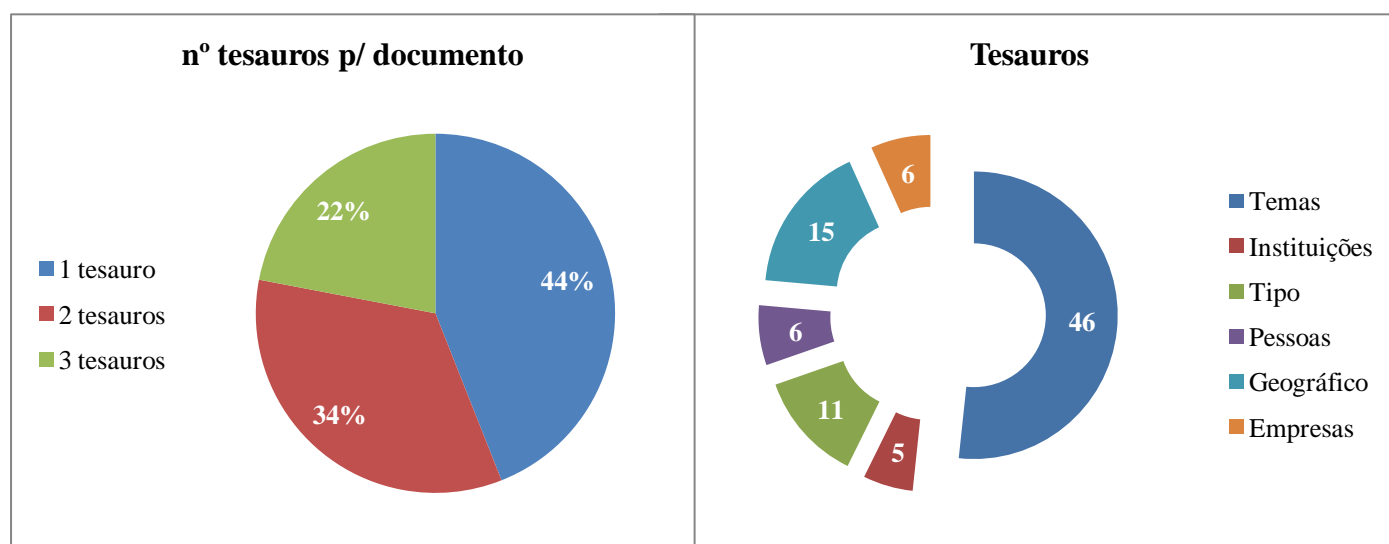


Gráfico 3 – Nº de Tesouros por documento

C- Recurso ao Infraconceito

Percebe-se que o recurso ao infraconceito é recorrente, principalmente nos conteúdos de opinião (artigos de opinião; crónicas ou editoriais). Os restantes vão sendo assinalados, embora de uma forma menos consistente e frequente. Desta forma, a tipologia textual é uma área temática a considerar na indexação e recuperação da informação, nomeadamente quando se pretende realçar a dimensão intelectual do conteúdo. Dos 50 documentos analisados, 11 foram contemplados com um descritor infraconceito.

Infra-conceito	OPINIÃO	PERFIL	REPORTAGEM	ENTREVISTA	INQUÉRITO	CRÓNICA
Nº documentos (total 11)	4	2	2	1	1	1

Tabela 13 – Aplicação de infraconceitos em tipologias texto diferenciadas

D- Harmonia do assunto com a secção onde o documento está inserido

As secções do jornal poderão influenciar, no momento da indexação, a escolha do descritor que irá representar o assunto do documento. Na verdade, a secção é a estrutura intelectual do jornal, funcionando como a primeira classificação dos conteúdos jornalísticos da imprensa. Por esse motivo, será de todo pertinente a adequação do descritor à secção em que o documento está publicado. Aliás, durante a análise dos tesouros, essa analogia foi considerada. E tal vai ao encontro do que foi analisado, pois verificou-se que dos 50 documentos, 29 foram indexados com descritores ajustados e relacionados com a secção correspondente, perfazendo 58% do total, o que confere o ajustamento natural da indexação com a secção. Observemos essa harmonia:

descritor	secção	documento
OPINIÃO	Opinião	1, 43
ECONOMIA PARALELA	DN Negócios	4
OPINIÃO	Editorial	5
AJUDA ÀS EMPRESAS	Economia e Negócios	9
SIMPLIFICAÇÃO LEGISLATIVA	Políticas e Conjuntura	10
POUPANÇA	DN Bolsa	11
CERTIFICADO DE AFORRO POUPANÇA	Dinheiro	12
RECESSÃO ECONÓMICA DESPEDIMENTO	Economia e Trabalho	14,15
CONTRATO DE TRABALHO FUNÇÃO PÚBLICA	Políticas e Conjuntura	17
ARTES	Artes	18
SALÁRIO	Economia	20
BANCO COMERCIAL PORTUGUÊS		
ABSENTISMO	Economia e Trabalho	24
TAXA DE DESEMPREGO	Economia e Trabalho	26
MONTEMOR-O-VELHO	País	27
LEIRIA	País	28
REFORMA ANTECIPADA	Economia	31
REFORMA DA ADMINISTRAÇÃO PÚBLICA	Economia	32
FUTEBOL	Desporto	34
VILA VERDE	Minho	36
HÓQUEI EM PARTINS	DN Sport	38
JOSÉ PESEIRO FUTEBOL	Desporto	39
BRAGA	Norte	41
AUDIÊNCIA	Televisão	45
JORNALISMO	Media e Televisão	46, 47
IMPRENSA	Media	48
ESTATUTO DO JORNALISTA	Televisão	49

Tabela 14 – Harmonia Assunto / Secção do jornal

E- Relação dos descritores com a linguagem natural do documento

Os descritores aplicados na indexação destes documentos estão relacionados com a estrutura interna, ou seja, com os títulos, legenda ou parágrafo-guia/*lead*. Dos 50 documentos analisados, é perceptível em 41 (82%) a proximidade existente entre o descritor e a linguagem natural presente em paratextos, havendo 9 (18%) onde tal não é visível, tendo sido necessário recorrer ao texto na íntegra para realizar a indexação. Dessa forma, depreende-se o cuidado em estreitar a ligação entre a linguagem controlada e o jargão jornalístico.

Observemos, a partir dos 41 documentos onde é visível essa relação, detetando 36 com descritores iguais ou semelhantes a palavras presentes nos títulos (72%); 32 no *lead* (64%) e 8 nas legendas (16%):

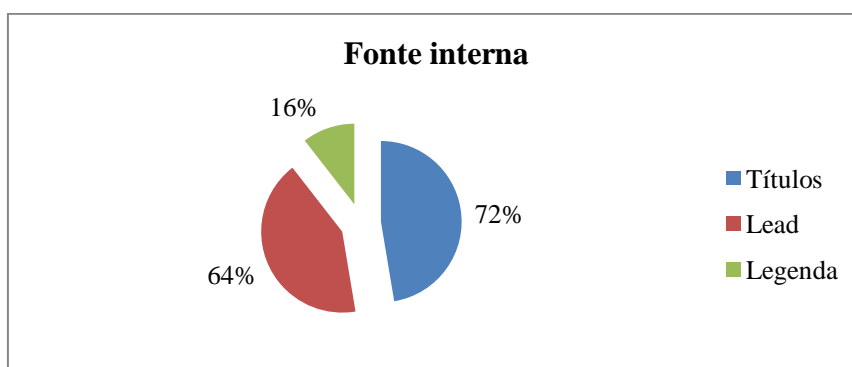


Gráfico 4 – Relação Descritores / Linguagem natural

3.2- Fotografias e infraconceitos

No novo sistema automatizado do Arquivo e Centro de Documentação do Grupo Controlinveste, todas as listas que constituem os Tesouros, já aqui referidos e analisados, sofreram revisões, tanto a nível estrutural como formal. No entanto, aqui apenas serão analisadas as alterações efetuadas no Tesouro 'Tipo' que, neste novo sistema, passou a ser denominado por "Infraconceitos". Isto porque os descritores que dele faziam parte serviam essencialmente essa função, ou seja, como ponto de acesso a diversas tipologias textuais. Contudo, percebe-se que foram incluídos novos descritores direcionados à indexação de fotografias. Relativamente a este aspeto, os termos *infraconceitos* já não estão estritamente ligados à forma do documento, mas mais a aspetos de conteúdo da imagem. E será nesse aspeto que a análise incidirá neste momento.



Esta é a lista atual com 82 descritores. Os que estão a azul e destacados a negrito são aqueles que foram inseridos neste novo sistema, que totalizam 45 novos termos.

ADEPTO	DECLARAÇÕES	MEIO CORPO
ADEREÇO	DIREITO DE	MODA
AGRESSÃO	RESPOSTA	NÚMEROS
ANÁLISE	DISCOGRAFIA	OBESO
AO TELEMÓVEL	EDITORIAL	OBITUÁRIO
ARTIGO DE	EFEMÉRIDES	OPINIÃO
OPINIÃO	ENTREVISTA	ORGANOGRAMA
BANDA DESENHADA	EQUIPA	PAISAGEM
BIBLIOGRAFIA	ESTATÍSTICA	PALMARÉS
BIOGRAFIA	FACHADA	PANORÂMICA
CARA	FARDADO	PERFIL
CASOS INSÓLITOS	FESTA	PLANTEL
CHORAR	FICHA	POBREZA
CLAQUE	FILMOGRAFIA	RECORTE
CLASSIFICAÇÃO	FOTO DE FAMÍLIA	SOCIAL
COM BOLA	FOTOLEGENDA	SORRIR
COMENTÁRIO	FRASE	TATUAGEM
CONFERÊNCIA DE	GENÉRICA	TIPO PASSE
IMPrensa	GESTO	TREINO
CORPO INTEIRO	GOLO	TRIBELA
CRONOLOGIA	GRANDE PLANO	TRISTE
CRONOLOGIA DA	GRÁFICO	TROFÉU
SEMANA	HISTÓRIA	TRÂNSITO
CRONOLOGIA DO	ÍCONE	VELHICE
MÊS	IDOSO	VISTA
CRÍTICA	INFOGRAFIA	VISTA AÉREA
CRÓNICA	INQUÉRITO	VISTA GERAL
CUMPRIMENTAR	INTERIOR	À CIVIL
DADOS	ISOLADO	
BIBLIOGRÁFICOS	MAPA	

Quanto à **forma**, verifica-se que alguns casos constituem desvios à NP 4036 (1992).

Vejamos:

uso de...	Descritores	NP 4036(1992)
VERBOS	Chorar Cumprimentar Sorrir	“ Verbos no infinitivo ou no participio não devem ser utilizados isoladamente como termos de indexação. As actividades devem ser representadas por nomes ou expressões verbais. (NP 4036, 1992: 11)



uso de...	Descritores	NP 4036(1992)
ADJETIVOS	Fardado Genérica Isolado Obeso Social Triste Grande plano Meio corpo	“Os adjetivos isolados podem ocorrer numa linguagem de indexação, nas circunstâncias especiais(...), mas o seu emprego dever sempre que possível evitado.” (NP 4036, 1992: 10)
PREPOSIÇÕES	Ao telemóvel Com bola À civil	As preposições são usadas nos termos compostos, servindo de elo de ligação entre o núcleo e o modificador. Por si só não constitui qualquer significado, já que existe como “combinação preposição nome (...) no termo composto” (NP 4036, 1992: 6), daí intercalar-se com nomes, sem que iniciem qualquer termo composto.

Tabela 15 – Infraconceitos: desvios à NP 4036 (1992)

Percebe-se que a escolha dos termos infraconceitos vai ao encontro da linguagem natural do jargão jornalístico, característica já denotada na análise dos Tesouros. O importante será fazer com que a aplicação dos infraconceitos seja útil na recuperação de aspetos técnicos e muito específicos dos conteúdos jornalísticos. Sem constrangimentos, no que diz respeito aos desvios das diretrizes da NP 4036 (1992), a criação deste tesauro impõe um carácter inovador no domínio das linguagens documentais. Claro está que se trata de algo consequente de uma abordagem específica. Por esse motivo, um tesauro desta tipologia terá de ser único, pois é construído tendo em conta as necessidades de recuperação de um meio de comunicação social, cuja informação e seus conteúdos circulam velozmente no circuito da informação.

Começemos por agrupar estes termos por **categorias**, para então perceber melhor a intencionalidade da utilização deste tesouro:

	Categorias	Descritores	%	
FOTOS	PESSOAS	adepto, claque, equipa, idoso, plantel	6%	59,8%
	OBJETOS	adereço, troféu	2,4%	
	ESTADOS EMOCIONAIS	chorar, sorrir, triste	3,7%	
	SITUAÇÕES COMPORTAMENTAIS	agressão, cumprimentar, gesto	3,7%	
	ASPETO FÍSICO	cara, corpo inteiro, meio corpo, obeso, velhice, tatuagem, fachada, fardado, interior, paisagem	12,2%	
	SITUAÇÕES SOCIAIS	ao telemóvel, casos insólitos, conferência de imprensa, efemérides, festa, moda, pobreza, social, trânsito, à civil, palmarés, treino, isolado	15,9%	
	ASPETOS TÉCNICOS	com bola, genérica, golo, grande plano, números, panorâmica, tribela, vista, vista aérea, vista geral, foto de família, tipo passe, ícone	15,9%	
TEXTOS	TIPOLOGIAS DOCUMENTAIS	análise, artigo de opinião, banda desenhada, bibliografia, biografia, comentário, cronologia, cronologia da semana, cronologia do mês, crítica, crónica, dados bibliográficos, declarações, direito de resposta, discografia, editorial, entrevista, estatística, ficha, filmografia, fotolegenda, frase, gráfico, infografia, inquérito, mapa, opinião, obituário, organograma, recorte, perfil, história, classificação	40,2%	

Tabela 16 – Categorias dos Infraconceitos

É perceptível que a categoria “Tipologias documentais” está vocacionada para indexar documentos da imprensa escrita, perfazendo 40,2% do total dos descritores. Partindo do pressuposto que as restantes categorias coadunam-se mais à indexação de fotografias, compreende-se que o tesouro infraconceitos, neste novo sistema integrado, foi reformulado no intuito de melhor recuperar aspetos importantes que as fotografias apresentam na construção dos conteúdos jornalísticos. Sendo assim, interessa recuperar informação sobre pessoas e seus estados emocionais, comportamentais e sociais, assim como o seu aspeto físico. Alguns objetos e seu aspeto também interessam e aspetos técnicos de vária ordem e da própria imagem. Todos estes aspetos perfazem 59,8% do total dos descritores deste tesouro.

Analisemos, agora, alguns exemplos:

Descritor	Foto/Descrição	Análise
ADEPTO	 <p>Povoia do Varzim , 02/06/2011: A selecção Nacional de Voleibol defrontou esta noite no Pavilhão Municipal da Póvoa do Varzim a Selecção Nacional da Argentina para a fase final da Liga Mundial. Adeptos portugueses.</p>	<p>Percebe-se que este descritor terá como função recuperar fotografias que possam ilustrar qualquer situação onde estejam presentes adeptos num ambiente desportivo. Esta tarefa genérica é também condicionada na descrição, onde se identifica na última frase a informação correspondente à utilidade contéudal que a foto poderá ter noutras utilizações. Por esse motivo, o descritor terá sido resultado de uma indexação automática, numa primeira fase.</p>
AO TELEMÓVEL	 <p>Telefone, telefones, telemovel, telemoveis, telecomunicacoes, rede movel, falar ao telemovel, operadoras, operadoras moveis.</p>	<p>Trata-se de um descritor que se aplica sempre que se pretenda recuperar uma situação social específica que, neste caso, é a imagem de alguém a falar ao telemóvel. Também se percebe pela descrição que se trata de um caso aleatório e não factual. Assim, o descritor terá sido resultado de uma indexação automática, numa primeira fase.</p>
CARA	 <p>Lisboa, 24/02/2012: hoje de manhã, na Federação Portuguesa de Futebol, Paulo Bento e Rui Jorge divulgam os convocados para os próximos jogos das Selecções. Paulo Bento responde aos jornalistas.</p>	<p>Todas as fotografias de personalidades, cuja cara surja em grande plano, são indexadas com este descritor. Verifica-se que não foi aplicada a indexação automática, pois o termo não se encontra no texto da descrição.</p>

Descritor	Foto/Descrição	Análise
COM BOLA	 <p>Prangins , 10/07/2012 - Segundo dia de treinos do FC Porto durante o estágio em Prangins, na Suíça, James Rodriguez</p>	<p>Este descritor, nitidamente resultante da linguagem natural, foi aplicado pelo conteúdo da imagem, entendendo-se a necessidade da recuperação num aspeto técnico, neste caso desportivo. Aqui não foi aplicada a indexação automática, pois o termo não se encontra no texto da descrição.</p>
GENÉRICA	 <p>Lisboa, 28/06/2011 - Reportagem : O turismo na costa alentejana. Praia da Azambuja.</p>	<p>O descritor não surge na descrição, daí não ser eleito através da indexação automática. Tratando-se de um conceito ambivalente, percebe-se que é aplicado com o intuito de mostrar a generalidade que a imagem pretende representar. Claro que só faz sentido este descritor quando associado a um outro temático, como é o caso com o termo “Praia” que também foi aplicado para esta imagem.</p>

Tabela 17 – Exemplos de aplicação de infraconceitos em fotografias



3.3- Linhas de atuação na indexação dos documentos analisados

Em sùmula, ficam as principais linhas de atuação das práticas de indexação aplicadas no Arquivo e Centro de Documentação do Grupo Controlinveste, linhas essas que vão ao encontro das variáveis analisadas nos diversos documentos. Como confirmação, foram aferidas as afirmações do Diretor do Arquivo e Centro de Documentação do Grupo Controlinveste dadas na entrevista cedida para este trabalho (cf. Anexo I, Vol. II), após a observação *in loco*.

1. Os documentos de conteúdo jornalístico da imprensa possuem uma estrutura muito específica, tendo o título, subtítulo, antetítulo ou títulos secundários, um papel crucial, já que possuem uma função apelativa para o leitor e, daí, condensar o tema da notícia. O parágrafo guia ou *lead* apresenta as ideias principais à volta das quais a notícia irá ser desenvolvida, respondendo às questões *o quê?*; *quem?*; *onde?* e *quando?*. Ora, estes serão uma das fontes onde o indexador irá “beber” no momento da indexação para de lá eleger os termos que representam o ou os assuntos do texto. Mas tal nem sempre se concretiza, principalmente quando o título é pouco claro ou, mesmo, abstrato. Isto acontece essencialmente em artigos de opinião e em entrevistas. Então, nestes casos, a análise de conteúdo do texto terá de ser realizada com mais primor. Tal corresponde ao que disse o Diretor do Arquivo e Centro de Documentação do Grupo Controlinveste na entrevista efetuada para este estudo “o que interessa é o conjunto do texto. Não é suficiente olhar para o título apenas. É preferível ler rapidamente um texto em diagonal do que estar a parar no título.” (cf. anexo I, Vol. II, pág.8).
2. Verifica-se uma tendência para o critério da especificidade na indexação dos conteúdos jornalísticos da imprensa, assim como a predominância na utilização de um tesouro, o temático. Contudo, é assumido não existir obrigatoriedade no número de descritores a aplicar e dos respetivos tesouros, assim como afirmou o Diretor do Arquivo e Centro de Documentação do Grupo Controlinveste na entrevista efetuada para este estudo “ Não há obrigatoriedade. (...) Tudo isto depende do bom senso dos documentalistas. Estamos a falar de um meio de comunicação... uma biblioteca de uma universidade, por exemplo, tem matérias



obrigatórias. Aqui não estamos a falar disso, mas sim informação que aparece agora e só depois surge anos depois.” (cf. anexo I, Vol. II, pág.8).

3. Outro aspeto interessante a salientar é a utilização de infraconceitos. Estes são descritores relativos à tipologia de documento e/ou abordagem, que servem para auxiliar a pesquisa. Num acervo com uma tipologia textual tão diversa, torna-se pertinente a procura por este parâmetro. Compreende-se a necessidade de recuperar informação só de artigos de opinião acerca de um determinado tema ou, mesmo, entrevistas a uma determinada personalidade. Muitos aspetos da vida social são medidos, daí a procura de estudos estatísticos resultantes de inquéritos realizados. Mas é nas fotografias que a sua aplicação se revela mais profícua, na medida em que conseguem recuperar aspetos específicos da imagem indexada, aspetos esses úteis na sua recuperação para futuras utilizações. Simões Dias, Diretor do do Arquivo e Centro de Documentação do Grupo Controlinveste, revela que “os infraconceitos foram criados porque nos dão jeito. Em qualquer texto sobre uma pessoa, interessa recuperar o que se diz sobre ela, dados que dizem respeito à biografia da pessoa, entrevistas em que a pessoa fala sobre algo ou frases da pessoa. Assim, a um nome de uma pessoa pode ser associada uma entrevista, uma frase, dados biográficos, etc. Isto no texto. Na fotografia começa a ser mais complexo, interessa recuperar determinados pormenores nas fotografias que para nós são importantes. Se precisarmos para uma notícia de um julgamento, da fachada de um tribunal, interessa procurar uma fachada especificamente. Há determinados pormenores que conseguimos recuperar mais facilmente ao indexar com infraconceitos, como por exemplo, corpo inteiro, cara, normalmente aspetos menos habituais mas que precisamos deles. Um que estamos a utilizar muito é o “genérica”. O que é uma foto genérica? É aquela que pode ser usada para circunstâncias semelhantes sem que identifique nada em concreto. Por exemplo, uma urgência de um hospital em que nada há que especifique qual é.” (cf. anexo I, Vol. II, pág.9).
4. Outro princípio a considerar no momento da indexação de conteúdos jornalísticos da imprensa relaciona-se, de alguma forma, com a estrutura física e intelectual do jornal, ou seja, com as suas secções. É óbvio que uma notícia

Maria José Rodrigues

estando integrada na secção ‘Economia’ do Jornal de Notícias terá, com toda a probabilidade, um descritor relacionado com esse assunto. No entanto, Simões Dias, Diretor do Arquivo e Centro de Documentação do Grupo Controlinveste, assume que a indexação não terá necessariamente de se relacionar com as secções dos jornais, argumentando que “Temos de olhar para isto num todo. Podemos ter o mesmo assunto que é tratado de uma forma mais generalista por um e de outra forma menos generalista por outro. É necessário, com a mesma indexação, tratar um assunto que é abordado no DN de uma forma lata e o no JN de uma forma menos lata. As secções dos jornais são tal e mais nada. Podem chamar o que lhes apetecer, por exemplo “Atual” não é nada. Como secção não vale nada, vale quando a informação é mapeada para o tesaurus. Por exemplo, hoje joga a seleção de futebol e para o jornal não é desporto, é atualidade, para o Arquivo é Desporto.”. Porém, onde se denota uma maior disparidade e diversidade de aplicação de termos será em notícias de secções mais abrangentes, como é o caso de ‘Sociedade’ ou ‘Norte-Sul’, devido à panóplia de assuntos tratados. Percebe-se, também, que uma notícia de âmbito regional e uma outra de âmbito nacional terão uma abordagem diferente no momento da indexação. Tratando-se de um arquivo de jornais nacionais, o processo de indexação terá de servir o que interesse para a redação e para a sua política editorial. Dessa forma, tudo o que tenha impacto e/ou interesse nacional, terá necessariamente uma indexação mais cuidada. Neste caso, poder-se-á recorrer a mais do que um descritor temático ou, mesmo, a outros descritores, sejam eles geográficos, de personalidades ou instituições. Como afirma Simões Dias, “Os jornais desportivos diferem dos mais generalistas, como o DN, enquanto que o JN existe a preocupação de indexá-lo por localidades, com o DN já é mais a parte temática. Com O Jogo, é mais a parte do Desporto e não o que se passa nesta ou naquela localidade.” (cf. anexo I, Vol. II, pág.7). O mesmo pode não ser verificado nas notícias de secções dedicadas a factos regionais. Neste estudo percebeu-se que nem sempre a mesma é indexada com precisão, sendo apenas referenciada com o descritor geográfico. Porém, resta a indefinição do que pode ser considerado de impacto nacional e regional. Que fatores os distinguem? Porém, o Diretor do Arquivo e Centro de Documentação do Grupo Controlinveste alega que o princípio não é esse quando lhe foi perguntado se as notícias de índole regional têm o mesmo tratamento no momento da indexação,

Maria José Rodrigues

dizendo “Têm de ter. Quando se pesquisa em Tesouros alguma coisa ligado ao Ambiente, tem de se ter textos genéricos e aqueles ligados a diferentes regiões e localidades.” (cf. anexo I, Vol. II, pág.8). Outra abordagem notória no momento da indexação de conteúdos jornalísticos da imprensa é reflexo do interesse do público pelas diversas modalidades desportivas. Isto porque uma notícia dedicada ao desporto-rei Futebol tem uma indexação mais completa do que outra relacionada com outro desporto. A uma notícia de futebol, são aplicados vários termos (tema, personalidades, equipas desportivas) enquanto que noutra de outro desporto apenas o termo relacionado com a própria modalidade. “É claro que não se indexa o nome de um jogador de um campeonato distrital, ao contrário do que acontece com os da 1ª e 2ª Liga.” (cf. anexo I, Vol. II, pág.9). Esta declaração de Simão Dias, Diretor do Arquivo e Centro de Documentação do Grupo Controlinveste na entrevista cedida para este estudo, vai ao encontro desta perspetiva.



CONCLUSÃO

Nesta dissertação de mestrado, conforme o que foi explanado na parte I, ficou bem assente o potencial informativo dos conteúdos jornalísticos da imprensa. Na verdade, eles suportam um valor de memória na medida em que transportam uma dimensão social e histórica. Também são portadores de uma narratividade específica, que influencia a forma como a informação é transmitida. Claro está que qualquer um destes aspetos é limitado pela democratização da divulgação desses mesmos conteúdos, na medida em que a política editorial tomará decisões que, necessariamente, irá marcar o cunho histórico e narrativo dos factos.

Compreende-se, nesta perspetiva, que a política de indexação desses mesmos conteúdos apresente características desse potencial. Com a observação realizada no Arquivo e Centro de Documentação do Grupo Controlinveste, como se pode constatar na parte II desta tese, o facto dos Tesouros existentes terem sido construídos a partir dos seus próprios conteúdos implica, necessariamente, que o valor de memória seja tido em conta. Claro está que as linhas de atuação no momento de indexação acentuarão, ou não, esse aspeto. O valor de legado que os arquivos e centros de documentação dos meios de comunicação social impõem tem de ser uma constante, assim como confirma Simões Dias, Diretor do Arquivo e Centro de Documentação do Grupo Controlinveste, quando afirma na sua entrevista que “este Arquivo é para ser usado hoje e amanhã” (cf. anexo I, Vol. II, pág.9). Pelo que foi constatado na parte II, no ponto 3, 3.1.1, chegou-se à conclusão que existe uma grande proximidade entre os descritores eleitos com a linguagem natural do conteúdo, bem visível no grau de especificidade dos termos usados. Para além disso, percebe-se, embora de uma forma involuntária, uma coerência entre a tradução de conceitos com a estrutura intelectual do jornal. A narratividade é também perceptível através da aplicação de infraconceitos nos conteúdos. Com eles, percebe-se a necessidade premente de extrair aspetos específicos de extrema importância no domínio da comunicação social. Tendo em conta que a velocidade da informação nestes meios é voraz, urge a implementação de formas cada vez mais céleres de recuperação de informação. Os arquivos e centros de documentação dos meios de comunicação social são verdadeiros intermediários do circuito da informação de um *media*. Os conteúdos produzidos nas redações são enviados para o arquivo, não para ficarem retidos em estado histórico ou definitivo, mas sim como disponíveis para uma nova utilização na



construção de um novo conteúdo. E foi nesta perspetiva que os infraconceitos foram criados e aplicados. Claro está que, à sua utilização, estão inerentes alguns desvios na forma dos descritores, conforme o que determina a NP 4036 (1992), como é exemplo a utilização de verbos, adjetivos e preposições. Então questiona-se: as orientações normativas podem contrariar os objetivos de um centro de documentação especializado? A resposta é dada na prática diária de qualquer arquivo ou centro de documentação. Contudo, se tal acontecer, dever-se-á, sem constrangimentos, contorná-la numa perspetiva de contributo para a mesma norma. Foi o que aconteceu no Arquivo e Centro de Documentação do Grupo Controlinveste com a criação dos já referidos infraconceitos, como é possível constatar neste trabalho, mais precisamente no ponto 3.2 da parte II.

Também se conclui que esse contributo é fruto de uma articulação conciliadora dos profissionais da informação, mais propriamente dos jornalistas e documentalistas. O público-alvo do Arquivo e Centro de Documentação do Grupo Controlinveste é mais abrangente do que propriamente os jornalistas das diferentes redações, como afirma o Diretor deste Arquivo na entrevista concedida para este estudo, quando diz que “o universo de leitores é tão grande que acabámos por alcançar um mercado inteiro sem grande dificuldade.” (cf. anexo I, Vol. II, pág.4). De facto, percebe-se que são os jornalistas o principal alvo na construção das linguagens documentais, já que existe uma grande proximidade ao jargão jornalístico. O circuito documental redação/arquivo já demonstra isso mesmo, quando os próprios conteúdos chegam ao centro de documentação com uma primeira indexação, realizada pelos jornalistas ou, então, determinada por uma indexação automática, confirmada depois pelo documentalista. Na entrevista cedida pelo Diretor deste Arquivo, é bem perceptível essa articulação, quando o mesmo afirma que “depois de estar no centro de documentação, estas tags, que são combinadas e escolhidas pela redação e que nos dão conhecimento, podem dar o nome que quiserem de acordo com o critério com que estão a escrever a notícia, serão depois mapeadas para os tesouros que utilizamos. Isto vai permitir que parte da informação, quando entra no arquivo, já esteja indexada. Esta indexação automática depois é confirmada pelos arquivistas.” (cf. anexo I, Vol. II, pág.4.).

Isto faz com que o mundo da ciência da informação seja menos hermético e mais transversal. E foi neste sentido que um dos objetivos desta tese foi atingido, na medida



em que é possível demonstrar de que forma a Ciência da Informação contribui na ação das Ciências da Comunicação, coexistindo uma relação de vizinhança entre as duas áreas.

Desta tese ressalta que esta relação acentua-se na nova era comunicacional em que vivemos. No mundo do digital, a massificação deu lugar a uma divergência informacional que os arquivos e centros de documentação dos meios de comunicação social têm de acompanhar. Para tal, sem dúvida que a indexação tem um papel preponderante, na medida em que consegue ser o centro do fluxo informacional, através das variantes da política de indexação aplicada. Será sempre no arquivo, no momento da indexação, que a canalização da informação encontra uma recuperação mais eficaz. Na era digital, o arquivo e centro de documentação de um meio de comunicação social tem uma função essencial que o pode transformar numa unidade de negócio, assim como afirma Simões Dias, Diretor do Arquivo e Centro de Documentação do Grupo Controlinveste “Isto não é para dar lucro, é para não dar prejuízo.” (cf. anexo I, Vol. II, pág.10).

Neste trabalho de investigação, foi pertinente e enriquecedor refletir sobre a forma como um Tesouro poderá retratar o perfil documental de um centro de documentação. Na verdade, se essa linguagem documental for construída numa metodologia indutiva, ou seja, a partir dos próprios conteúdos, tal acontecerá. “O tesouro é uma ferramenta que serve para filtrar informação” (cf. anexo I, Vol. II, pág.7), assim afirma o Diretor do Arquivo e Centro de Documentação do Grupo Controlinveste na entrevista concedida e, de facto, ele tanto filtra como desenha as temáticas tratadas num determinado meio de comunicação. Se esse filtro for realizado com ponderação por parte do documentalista, esta linguagem documental terá outras funções que não só a do “filtro”.

Por fim, ficam sugestões para futuros trabalhos nesta área. Cruzar a Ciência da Informação com a Ciência da Comunicação proporciona uma enorme variedade de temáticas a serem exploradas em estudos académicos. Este trabalho assentou na observação de conteúdos jornalísticos da imprensa. Porém, o mundo da imagem facilita inúmeros aspetos a investigar. Aqui, a fotografia foi levemente explorada a propósito de um aspeto inovador na construção das linguagens documentais. Porém, um estudo mais aprofundado sobre a indexação de fotografias com fins jornalísticos será, certamente



Maria José Rodrigues

uma mais-valia para o desenvolvimento científico destas áreas. O mesmo acontece com a indexação do vídeo em uso jornalístico, cada vez mais solicitado na nova era comunicacional e digital.

Sem dúvida, que a Indexação de conteúdos jornalísticos apresenta um mundo informacional ainda a explorar. Urge conhecer os trilhos da informação desvendados pela política de indexação. Esta deve abrir portas e nunca fechá-las. O intuito desta tese foi apenas abrir uma janela.



REFERÊNCIAS

1. Normas

NP 3715 (1989). *Documentação: Método para a análise de documentos, determinação do seu conteúdo e seleção de termos de indexação*. Lisboa: Instituto Português da Qualidade.

NP 4036 (1992). *Documentação: Tesouros monolíngues: directivas para a sua construção e desenvolvimento*. Lisboa: Instituto Português da Qualidade.

2. Imprensa

Abrantes, José C. (2005). *As fontes anónimas no DN*. Diário de Notícias, 25/05/2005.

Bernardino, Carla (2012). *Computadores substituem jornalistas*. Notícias TV, 06/07/2012.

Bernardino, Carla (2012). *Como a RTP guarda a memória do país*. Notícias TV, 30/03/2012.

Campos, Anabela; Coelho, Liliana (2012). *Balsemão aguarda reação do Governo contra a Google*. Expresso, 07/11/2012 [Em linha]. Disponível em <http://expresso.sapo.pt/gen.pl?p=print&op=view&fokey=ex.stories/765102&sid=ex.s>. [Consult. 07 novembro 2012].

Cardoso, Joana A. (2005). *Novos media estão a mudar os comportamentos colectivos*. Público, 09/10/2005.

Dossier Media: Entrevista com Ricardo Costa, director adjunto de Informação na SIC Notícias (2006). Público, 03/06/2006.

Jornais brasileiros impedem Google usar notícias de borla (2012). Expresso, 20/10/2012 [Em linha]. Disponível em <http://expresso.sapo.pt/gen.pl?p=print&op=view&fokey=ex.stories/761222&sid=ex.s>. [Consult. 21 outubro 2012].

Madeira, Paulo M. (2006). *Jornal americano convida leitores a escolher primeira página*. Público, 29/01/2006.

Mota, Sérgio B. (2007). *Governo de Lula critica poder dos media*. Diário de Notícias, 05/02/2007.

O futuro da imprensa está na internet paga (2012). Diário de Notícias, 25/05/2012.

Pinheiro, Susana (2005). *Peritos dizem que futuro do jornalismo é 'online'*. Diário de Notícias, 04/06/2005.

Silveira, Ana Filipe (2012). *'Depois do Adeus? Vai passar pelo DN'*. Notícias TV, 30/03/2012.

Torres, Eduardo C. (2005). *Cidadãos quê?*. Público, 01/11/2005.

3. Estudos

Araújo, Eliany Alvarenga (2002). *O fenômeno informacional na ciência da informação: abordagem teórica-conceitual*. In *Ciência da Informação e Biblioteconomia: múltiplos discursos*, cap.1. São Luís do Maranhão, Edufma; Edfama.

Araújo, Raquel Leite (2010). *Indexação em bibliotecas jurídicas: um estudo exploratório*. [Em linha]. [Consult. 09 mar. 2012]. Disponível em: <http://hdl.handle.net/10183/27807>.

Araujo Junior, Rogério Henrique de; Tarapanoff, Kira (2006). *Precisão no processo de busca e recuperação da informação: uso da mineração de textos*. *Ciência da informação*. [Em linha], Disponível em <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0100-19652006000300023&lng=en&nrm=iso>. [Consult. 01 de mar. 2012].

Barbosa, Susana (s.d.). *Bases de dados e webjornalismo: em busca de novos conceitos*. Universidade Federal da Bahia [Em linha] Disponível em <http://www.bocc.ubi.pt/pag/barbosa-suzana-bases-dados-webjornalismo.pdf>> [Consult. 25 março 2012].

Borges, Graciane Silva Bruzinga (2009). *Indexação automática de documentos textuais: critérios essenciais*. [Em linha]. [Consult. 17 fev. 2012]. Disponível em: <http://www.bibliotecadigital.ufmg.br/dspace/handle/1843/ECID-7W5JH9>.

Borko, Harold (1968). *Information Science – what is it?* Washington, American documentation.

Café, Lígia; Brascher, Marisa. (2008). *Organização da informação e bibliometria*. *Encontros Bibli: revista eletrônica de biblioteconomia e ciência da informação* [Em linha]. [Consult. 10 jan. 2012]. Disponível em: <http://www.periodicos.ufsc.br/index.php/eb/article/view/1334/1032>.

Canavilhas, João Messias (s.d.). *Webjornalismo: considerações gerais sobre jornalismo na web*. Universidade da Beira Interior. [Em linha] Disponível em <http://www.bocc.ubi.pt/pag/canavilhas-joao-webjornal.pdf>> [Consult. 25 de março 2012].

Cardoso, Gustavo; Espanha, Rita; Araújo, Vera (2009). *Da comunicação de massa à comunicação em rede*. Coleção Comunicação, vol. 18. Porto, Porto Editora.

Castro, César Augusto, org. (2002). *Ciência da Informação e Biblioteconomia: múltiplos discursos*. São Luís do Maranhão, Edufma, Edfama.



- Cavalcanti, C.R. (1978). *Indexação e Tesouro: metodologia e técnica*. Brasília, ABDF.
- Correia, João C. (2006). *Do texto ao contexto: o triângulo “sistema”, “esfera pública” e “mundo da vida”*. Jornadas Internacionais de Jornalismo. 2006. Horizontes do jornalismo. Porto, Universidade Fernando Pessoa.
- Costa, Maria Leonor Lopes Fantesia Pereira (2009). *Definição de uma Política de Indexação numa Biblioteca Escolar e a Recuperação da Informação*. Dissertação de Mestrado. Faculdade de Letras da Universidade de Lisboa.
- Currás, Emilia (1998). *Tesauros: manual de construcción y uso*. Madrid, Kaher.
- Dalmonete, Edson Fernando (2010). *Presente: o tempo do jornalismo e seus desdobramentos*. História, vol.29, n.1, 2010. [Em linha] Disponível em <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0101-90742010000100019&lng=en&nrm=iso>. ISSN 1980-4369. <http://dx.doi.org/10.1590/S0101-90742010000100019>. [Consult. 19 de abril 2012].
- Fernandes, José Manuel (2011). *Liberdade e Informação*. Coleção Ensaio da Fundação, vol.21. Lisboa, Fundação Francisco Manuel dos Santos.
- Fidalgo, António (s.d.). *Um exame do jornalismo online*. Recensão. Universidade da Beira Interior. [Em linha] Disponível em <http://www.recensio.ubi.pt/modelos/recensoes/recensao.php3?codrec=42> [Consult. 25 março 2012].
- Fidalgo, Joaquim (2008). *O jornalista em construção*. Coleção Comunicação, vol.17. Porto, Porto Editora.
- Freitas, Judite A. Gonçalves, dir.; Gouveia, Luís Borges; Regedor, António Borges, eds. (2012). *Ciência da Informação: contributos para o seu estudo*. Porto, Universidade Fernando Pessoa.
- Fujita, Mariângela Spotti Lopes; Rubi, Milena Polsinelli. (2006). *O ensino de procedimentos de política de indexação na perspectiva do conhecimento organizacional: uma proposta de programa para a educação à distância do bibliotecário*. Perspectivas em Ciência da Informação, 11(1), 48-67. [Em linha]. [Consult. 23 fev. 2012]. Disponível em: http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1413-99362006000100005&lng=en&tlng=pt. <http://dx.doi.org/10.1590/S1413-99362006000100005>.
- Izquierdo Arroyo, José Maria; Moreno Fernández, Luís Miguel (s.d.). *Listas de encabezamientos de matéria y thesauri en perspectiva comparada*. Murcia, Universidad de Murcia.
- Jornadas Internacionais de Jornalismo. 2006. Horizontes do jornalismo: atas, Porto, Universidade Fernando Pessoa.

Maria José Rodrigues

Lapa, Remi C.; Corrêa, Renato F. (2011). *Leis bibliométricas aplicadas na indexação automática de testes e dissertações da UFPE*. A gestão da informação na era do conhecimento [Em linha]. [Consult. 23 jan. 2012]. Disponível em: www.repositorios.ufpe.br/jspui/bitstream/123456789/85/.../trab35.pdf. ISBN978-85-60323-31-9.

Lancaster, Frederick Wilfrid (2002). *El control del vocabulario en la recuperación de información*. Valencia, Universitat de Valencia.

Lancaster, Frederick Wilfrid (2004) *Indexação e resumos: teoria e prática*. 2. ed. Brasília: Briquet de Lemos.

Lima, Vania Mara Alves; Boccato, Vera Regina Casari (2009). *O desempenho terminológico dos descritores em Ciência da Informação do Vocabulário Controlado do SIBi/USP nos processos de indexação manual, automática e semi-automática*. *Perspect. ciênc. inf.* [Em linha]. [Consult. 23 jan. 2012]. Disponível em: http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1413-99362009000100010&lng=en&nrm=iso.

Lucas, Clarinda Rodrigues (1996). *Indexação: gosto de leitura do bibliotecário*. [Em linha]. [Consult. 09 mar. 2012]. Disponível em: <http://www.bibliotecadigital.unicamp.br/cgi-bin/search.cgi?>

Mamfrim, Flávia P.B. (1991). *Representação de conteúdo via indexação automática em textos integrais em língua portuguesa*. *Ciência da Informação* [Em linha]. [Consult. 23 jan. 2012]. Disponível em: <http://revista.ibict.br/index.php/ciinf/article/view/1261/899>.

Neveu, Érik (2005). *Sociologia do jornalismo*. Coleção Comunicação, vol.5. Porto, Porto Editora.

Pimenta, Márcia Teresa da Rocha (2002). *O profissional da informação e as novas mediações no atendimento ao cliente*. In *Ciência da Informação e Biblioteconomia: múltiplos discursos*, cap.1. São Luís do Maranhão, Edufma; Edfama.

Pinto, Manuel; Marinho, Sandra (2003). *A qualidade em jornalismo: problematização e operacionalização do conceito*. [Em linha]. Disponível em <http://hdl.handle.net/1822/997>. [Consult. 18 abril 2012].

Pinto, Ricardo J.; Sousa, Jorge P. (1998). *O futuro incerto da internet: intercomunicar além do comércio e da publicidade*. [Em linha]. Disponível em bocc.ubi.pt/pag/sousa-jorge-pedro-futuro_net.pdf. [Consult. 10 julho 2012].

PORTUGAL. Biblioteca Nacional. Área de Classificação e Indexação (1998). *Siporbase: Sistema de Indexação em português: manual*. Lisboa, Biblioteca Nacional.

Ribeiro, Fernanda (1996). *Indexação e controlo de autoridade em arquivos*. [Em linha]. [Consult. 23 jan. 2012]. Disponível em: <http://repositorio-aberto.up.pt/handle/10216/10721>.



Rieffel, Rémy (2003). *Sociologia dos media*. Coleção Comunicação, vol.3. Porto, Porto Editora.

Robredo, Jaime (1991). *Indexação automática de textos: uma abordagem otimizada e simples*. Ciência da Informação, Brasília, DF, Brasil, 20, ago. 1991. [Em linha]. [Consult. 17 fev. 2012]. Disponível em: <http://revista.ibict.br/ciinf/index.php/ciinf/article/view/1349/976>.

Serra, Paulo J. (2007). *Manual de Teoria da Comunicação*. Universidade da Beira Interior. [Em linha]. Disponível em www.livroslabcom.ubi.pt/book/51. [Consult. 10 de julho 2012].

Silva, Armando Malheiro; Ribeiro, Fernanda (2002). *Das «ciências» documentais à ciência da informação: ensaio epistemológico para um novo modelo curricular*. Coleção Biblioteca das Ciências do Homem. Porto, Edições Afrontamento.

Slype, Georges van (1991). *Los lenguajes de indización: concepción, construcción y utilización en los sistemas documentales*. Salamanca, Fundación Germain Sánchez Ruy Pérez.

Strehl, Leticia (1998). *Avaliação da consistência da indexação realizada em uma biblioteca universitária de artes*. [Em linha]. [Consult. 23 jan. 2012]. Disponível em: http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0100-19651998000300011#back.

Waller, Suzanne; Masse, Claudine [colab.] (1999) – *L'analyse documentaire: une approche méthodologique*. Paris: ADBS Editions.

Wolf, Mauro (1999). *Teorias da Comunicação*. Lisboa, Editorial Presença.

[em branco]