

Curso de Bacharelado em Biblioteconomia na Modalidade a Distância

Edivanio Duarte de Souza

Análise da Informação

Semestre

3



UNIVERSIDADE FEDERAL DO RIO DE JANEIRO



1920 | 2020

Rio de Janeiro

Faculdade de Administração
e Ciências Contábeis
Departamento
de Biblioteconomia

2018



Permite que outros remixem, adaptem e criem a partir do seu trabalho para fins não comerciais, desde que atribuam o devido crédito ao autor e que licenciem as novas criações sob termos idênticos.

Presidência da República

Ministério da Educação

Coordenação de Aperfeiçoamento de Pessoal de Nível

Superior (CAPES)

Diretoria de Educação a Distância (DED)

Sistema Universidade Aberta do Brasil (UAB)

Universidade Federal do Rio de Janeiro (UFRJ)

Núcleo de Educação a Distância (NEAD)

Faculdade de Administração e Ciências Contábeis (FACC)

Departamento de Biblioteconomia

Leitor

Vânia Guedes

Comissão Técnica

Célia Regina Simonetti Barbalho

Helen Beatriz Frota Rozados

Henriette Ferreira Gomes

Marta Lígia Pomim Valentim

Comissão de Gerenciamento

Mariza Russo (in memoriam)

Ana Maria Ferreira de Carvalho

Maria José Veloso da Costa Santos

Nadir Ferreira Alves

Nysia Oliveira de Sá

Equipe de apoio

Eliana Taborda Garcia Santos

José Antonio Gameiro Salles

Maria Cristina Paiva

Miriam Ferreira Freire Dias

Rômulo Magnus de Melo

Solange de Souza Alves da Silva

Coordenação de

Desenvolvimento Instrucional

Cristine Costa Barreto

Desenvolvimento instrucional

Renata Vittoretti

Diagramação

Patrícia Seabra

Revisão de língua portuguesa

Beatriz Fontes

Projeto gráfico e capa

André Guimarães de Souza

Patrícia Seabra

Normalização

Dox Gestão da Informação

S719a Souza, Edivanio Duarte de.

Análise da informação / Edivanio Duarte de Souza; [leitora] Vânia Lisbôa da Silveira Guedes. – Brasília, DF: CAPES: UAB; Rio de Janeiro, RJ: Departamento de Biblioteconomia, FACC/UFRJ, 2018.

98p.: il.

Inclui bibliografia.

ISBN 978-85-85229-60-3 (brochura)

ISBN 978-85-85229-61-0 (e-book)

1. Análise semântica. 2. Serviços de informação. I. Guedes, Vânia Lisbôa da Silveira. II. Título.

CDD 025.4

DU 025.4

Catálogo na publicação por: Miriam Dias CRB-7 / 6995

Caro leitor,

A licença CC-BY-NC-AS, adotada pela UAB para os materiais didáticos do Projeto BibEaD, permite que outros remixem, adaptem e criem a partir desses materiais para fins não comerciais, desde que lhes atribuam o devido crédito e que licenciem as novas criações sob termos idênticos. No interesse da excelência dos materiais didáticos que compõem o Curso Nacional de Biblioteconomia na modalidade a distância, foram empreendidos esforços de dezenas de autores de todas as regiões do Brasil, além de outros profissionais especialistas, a fim de minimizar inconsistências e possíveis incorreções. Nesse sentido, asseguramos que serão bem recebidas sugestões de ajustes, de correções e de atualizações, caso seja identificada a necessidade destes pelos usuários do material ora apresentado.



LISTA DE FIGURAS

Figura 1 – Processo de representação da informação.....	19
Figura 2 – Representação descritiva (ficha catalográfica – AACR2).....	20
Figura 3 – Representação descritiva (catálogo eletrônico)	21
Figura 4 – Representação descritiva (Referência – NBR 6.023/ABNT).....	22
Figura 5 – Medidas de extensão da indexação de um documento	24
Figura 6 – Representação da informação no sistema de recuperação da informação	27
Figura 7 – Análise da informação no sistema de recuperação da informação	28
Figura 8 – Coeficiente de precisão	29
Figura 9 – Coeficiente de revocação.....	30
Figura 10 – Esquema geral de estratégia de leitura.....	48
Figura 11 – Esquemas da CDD 22ª edição.....	52
Figura 12 – Tabelas auxiliares da CDD 22ª edição	53
Figura 13 – Lista de cabeçalhos de assunto	55
Figura 14 – Estrutura de relações semânticas no tesouro	58
Figura 15 – Índice alfabético de assunto.....	62
Figura 16 – Índice e resumo: materialidade e formalidade	63
Figura 17 – Literatura concisa	65
Figura 18 – Minirresumo.....	65
Figura 19 – Resumo estruturado	65
Figura 20 – Resumo modular	66
Figura 21 – Resumo telegráfico.....	68
Figura 22 – Esquema de elaboração de resumos	68
Figura 23 – Resumo altamente formatado	69

LISTA DE QUADROS

Quadro 1 – Tipos de representação temática.....	45
Quadro 2 – Tipos de representação temática.....	45
Quadro 3 – Classificações bibliográficas	52
Quadro 4 – Análise e síntese do conteúdo do documento.....	53
Quadro 5 – Análise e síntese do conteúdo usando a CDD	54
Quadro 6 – Categorias de análise lógica.....	87
Quadro 7 – Estrutura de análise diplomática de um livro	89



SUMÁRIO

	APRESENTAÇÃO DA DISCIPLINA	9
1	UNIDADE 1: FUNDAMENTOS DA ANÁLISE DA INFORMAÇÃO	13
1.1	INTRODUÇÃO.....	15
1.2	REPRESENTAÇÃO DA INFORMAÇÃO.....	18
1.2.1	Representação descritiva da informação.....	20
1.2.2	Representação temática da informação.....	22
1.3	ORGANIZAÇÃO DA INFORMAÇÃO.....	25
1.4	RECUPERAÇÃO DA INFORMAÇÃO.....	28
	RESUMO	31
1.5	Atividade.....	32
2	UNIDADE 2: PROCEDIMENTOS DE ANÁLISE DA INFORMAÇÃO	35
2.1	LINGUAGEM NATURAL E LINGUAGEM ARTIFICIAL.....	37
2.2	ESTRATÉGIAS DE LEITURA.....	47
2.3	INSTRUMENTOS AUXILIARES E PRODUTOS.....	51
2.4	PRÁTICA DA ANÁLISE DA INFORMAÇÃO.....	70
	RESUMO	73
2.5	Atividade.....	74
3	UNIDADE 3: INTERDISCIPLINARIDADE NA ANÁLISE DA INFORMAÇÃO	79
3.1	CONTRIBUIÇÕES DA LINGÜÍSTICA E DA TERMINOLOGIA.....	82
3.2	CONTRIBUIÇÕES DA LÓGICA.....	86
3.3	CONTRIBUIÇÕES DA DIPLOMÁTICA.....	88
	RESUMO	91
3.4	Atividade.....	92
	REFERÊNCIAS	94



APRESENTAÇÃO DA DISCIPLINA

Os diferentes sistemas de informação, incluindo a biblioteca, o centro de documentação e o centro de informação, entre outros, são planejados objetivando a recuperação da informação. Para tanto, faz-se necessário primeiramente tratar e organizar os conteúdos informacionais de forma a possibilitar a busca e a localização da informação desejada, no momento oportuno. A organização da informação é condicionada pelo tratamento dos conteúdos dos documentos e, especificamente, pelo estabelecimento de informações documentárias, que possibilitam a busca e aperfeiçoam a recuperação da informação.

Destaca-se nesse domínio a representação da informação, que é executada a partir dos processos de análise e síntese dos aspectos físicos e do conteúdo informativo do documento. Nesse contexto, a análise da informação é uma modalidade de leitura desses elementos, que toma como referência o próprio documento e as condições de produção e de uso, que são dele constitutivas. Assim, é compreendida a partir de uma perspectiva pragmática em que o documento é considerado em sua dinâmica como constituído de informações com possibilidade de diferentes leituras e, por conseguinte, distintas formas de análises e de produção de informações documentárias. A síntese, por sua vez, é realizada a partir das relações entre os diferentes elementos representativos das unidades de informação que compõem os documentos, conformando-se em informação documentária.

A disciplina *Análise da Informação* é o primeiro componente do **Eixo 2: Organização e Representação da Informação** da matriz curricular do **Curso de Graduação em Biblioteconomia na modalidade a distância**. De modo geral, esta disciplina se apresenta como uma introdução aos estudos sobre tratamento e organização da informação, uma vez que fundamenta o planejamento, a elaboração, o uso e a avaliação dos diferentes instrumentos, processos e produtos, nesta subárea de conhecimento. Trata-se, portanto, de uma atividade que compõe uma das principais habilidades do bibliotecário e, por conseguinte, encontra-se na centralidade da formação deste profissional. A rigor, a análise da informação é a base para a realização de uma série de processos documentários, tais como catalogação, classificação, indexação e elaboração de resumos, além de outras aplicações, como, por exemplo, a definição de área de cobertura da linguagem documentária e as análises de perguntas na elaboração de estratégias de busca.

A presente disciplina tem como objetivo introduzir o aluno de Biblioteconomia na prática analítico-sintética que fundamenta a organização e a representação da informação. E, nesse horizonte, pretende, especificamente, discutir os principais fundamentos da análise da informação nos domínios da representação, organização e recuperação da informação; apresentar as condições e o conjunto de procedimentos que compõem o processo da análise da informação e discutir os principais elementos que subsidiam os processos de análise e representação da informação no domínio interdisciplinar da Linguística, Terminologia, Lógica e Diplomática.

Na Unidade 1, **Fundamentos da Análise da Informação**, observa-se que o domínio pragmático da análise da informação não se restringe à representação do conteúdo do documento, mas alcança diversas atividades que compõem a representação, a orga-

nização e a recuperação da informação, tais como a construção e o uso de instrumentos de controle de vocabulário e a composição de produtos e serviços informacionais.

Complementarmente, situa a análise da informação no domínio conceitual plural, que decorre da dupla influência presente na produção científica, europeia e norte-americana, advindas, respectivamente, da Documentação e da Biblioteconomia. Além dessas, outras áreas, como Ciência da Informação e Recuperação da Informação, relacionam-se a esse domínio de estudo. Em que pese a dispersão terminológica que caracteriza esta subárea, considera-se como mais adequada a expressão análise da informação, que, ao buscar alcançar o conteúdo informativo, contempla os aspectos extrínsecos (descritivos) e intrínsecos (temáticos) do documento.

Na Unidade 2, **Procedimentos de Análise da Informação**, observa-se que o estudo da análise da informação deve ser realizado no ponto de integração entre a linguagem natural, a linguagem especializada e a linguagem documentária. A primeira se refere à linguagem utilizada pelas diferentes pessoas no contexto social mais amplo. A segunda, às diferentes linguagens utilizadas nas diversas áreas do conhecimento. E a terceira, que é elemento central do processo de tratamento e organização da informação, compõe o conjunto de linguagens formais adotadas na construção de informações documentárias. A linguagem documentária é, pois, uma metalinguagem porque tradutora das outras, no interior dos sistemas de informação. Ela, em última análise, tem como finalidade estabelecer a comunicação entre os diferentes universos informacionais que compõem um campo semântico.

A análise da informação é um processo complexo que envolve um conjunto de estratégias, procedimentos e técnicas destinados a identificar e expressar o conteúdo informativo dos documentos por intermédio de informações documentárias. Este processo tem como objetivo a representação e a organização do conhecimento registrado e como finalidade a recuperação da informação. O tratamento da informação contempla uma redução semântica realizada a partir tradução do conteúdo informativo do documento em informação documentária. Para tanto, adota estratégias de leitura que permitem a integração de processos dinâmicos de interpretação, como analogias, inferências e predições.

O aprimoramento da prática leitora é realizado a partir do estabelecimento de estratégias de leitura que devem fazer parte do rol de habilidades do profissional bibliotecário. Com efeito, a leitura é um processo interativo dinâmico que pode apresentar níveis diferenciados de complexidade, dependendo da relação entre o texto/documento e o leitor/analista. A rigor, a leitura compreende um universo de referência que condiciona a formação de esquemas conceituais e o conhecimento prévio que possibilita a análise das diferentes unidades de informação que compõem os documentos. Em outros termos, a leitura depende do conhecimento prévio sobre o conteúdo em análise que compõe a condições imediatas de cognição, interpretação e análise dos conteúdos informacionais.

Na prática, a análise da informação objetiva o exame dos conteúdos dos documentos e a identificação de unidades de informação deles representativas. Essas unidades de informação são estruturas significantes que servem de base para construção das informações documentárias. De modo mais específico, tem como objetivo a leitura, a identificação, a seleção e a descrição de unidades linguísticas que formam os conteúdos informacionais.

A informação documentária é, em sua essência, analítico-sintética e, por conseguinte, indicativa do conteúdo dos documentos a que se refere. A construção dessa categoria de informação é um dos principais fundamentos para organização e recuperação do conjunto de informações registradas nos documentos que formam uma coleção. Ela

incide diretamente no conteúdo informacional, nos procedimentos, nos instrumentos e nos produtos que formam o conjunto de processos de tratamento e organização da informação.

Na Unidade 3, **Interdisciplinaridade na Análise da Informação**, faz-se necessário compreender a análise da informação na relação com diferentes processos de tratamento e organização da informação, que implica entendimento da relação da prática analítica com a informação contida nos documentos e nas informações a partir delas produzidas. Além disso, observar-se-á que a análise da informação não se restringe à estrutura textual, uma vez que não se dedica apenas à frase, mas ao texto e ao contexto, exigindo uma abordagem interdisciplinar por intermédio da colaboração de diferentes disciplinas, como *Diplomática, Linguística, Lógica e Terminologia*, entre outras.

A elaboração e o uso de linguagens documentárias compreendem o domínio da área do conhecimento em que se insere e do campo de referência em que se encontram os usuários do sistema de informação em que serão adotadas. O fato é que a polissemia que caracteriza a constituição dos acervos documentais possibilita diferentes leituras, apontando para a necessidade do estabelecimento de quadros de referências na definição de informações documentárias que permitem a recuperação dos conteúdos informacionais. A análise com base interdisciplinar é essencial para a determinação formal de conceitos, bem como das relações semânticas e sintáticas, na estruturação e no uso das linguagens documentárias.



UNIDADE 1

FUNDAMENTOS DA ANÁLISE DA INFORMAÇÃO





1.1 INTRODUÇÃO

O planejamento e a organização de sistema de informação têm por base dois princípios complementares, a saber, o princípio analítico e o princípio sintético. Com efeito, a análise e a síntese da informação estão presentes desde a definição dos itens que compõem o acervo da biblioteca ao efetivo uso de informações. Contudo, a análise da informação é abordada, principalmente, nos domínios dos processos de representação e recuperação da informação.

O principal fundamento da análise da informação concerne ao processo de intermediação que caracteriza os sistemas de informação, sobretudo em função da impossibilidade de acesso direto ao universo de documentos. Em que pesem as diferentes possibilidades de acesso à informação de textos completos, especialmente nos espaços digitais, não se pode perder de vista a importância fundamental da representação nos processos de organização e recuperação da informação. Muitas vezes, não se trata apenas de aperfeiçoar o acesso ou promover a celeridade deste, mas efetivamente de dar condições aos diferentes tipos de usuários à recuperação das informações de que necessitam. Assim, a presente unidade tem como objetivo situar a disciplina **Análise da Informação** nos domínios da **representação, organização e recuperação da informação**.

É importante destacar que uma abordagem adequada para tratar a análise da informação diz respeito àquela que tem como base a *Teoria de Sistema de Ludwig Von Bertalanffy*. Trata-se de compreender os diversos tipos de centros de informação em uma concepção sistêmica, ampla e usualmente, denominada de sistema de recuperação da informação.

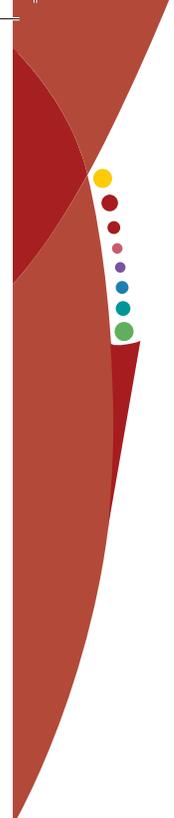


Explicativo

Teoria de Sistema

A *Teoria Geral de Sistema* foi desenvolvida pelo biólogo austríaco, Ludwig von Bertalanffy (1901-1972), na década de 1930. A partir da crítica ao cartesianismo, que se caracterizava pelo reducionismo analítico, este estudioso defendeu a necessidade da adoção de uma visão orgânica da biologia na compreensão e explicação de diversos sistemas. Com base nessa teoria, dever-se-ia considerar a globalidade dos organismos e a interdependência entre as suas partes constituintes. Com efeito, quando tratadas de forma integrada, as partes que constituem a unidade orgânica apresentam qualidades e/ou quantidades distintas daquelas ao serem tratadas isoladamente. A ideia central da sua teoria é que qualquer unidade orgânica é maior que a soma de suas respectivas partes. Essa teoria foi usada amplamente a partir dos anos 1950, sobretudo, no campo da cibernética. Um sistema, portanto, corresponde a um conjunto de partes integradas que se organizam e funcionam, de forma interdependente, em torno de uma finalidade comum.

Fonte: BERTALANFFY, L. V. **Teoria geral dos sistemas: fundamentos, desenvolvimentos e aplicações**. Petrópolis, RJ: Vozes, 2008. 360p.



Inicialmente é importante esclarecer os fundamentos e o domínio conceitual em que se estabelece a análise da informação, sobretudo porque, além da importância dessa atividade no sistema de informação, a produção científica acerca dessa temática é marcada pelo pluralismo conceitual. Nesse domínio, você vai encontrar, pelo menos, três expressões distintas que são marcadas pelo hibridismo conceitual e pela sobreposição de alguns elementos: análise documentária, análise de assunto e análise da informação.

Para melhor compreensão, deve-se considerar que essas expressões têm origens e fundamentos diversos, porém muito próximos. Assim, não se trata apenas de formas diferentes de expressar a mesma temática ou a mesma atividade no domínio de um sistema de recuperação da informação. A rigor, essas concepções têm por base dois modelos de abordagens historicamente distintos, o europeu e o norte-americano, que têm por fundamentos, respectivamente, a Documentação e a Biblioteconomia.

A expressão “análise documentária”, de origem francesa (*analyse documentaire*), vem do campo da Documentação, que “[...] é definida como um conjunto de procedimentos efetuados com o fim de expressar o conteúdo de documentos, sob formas destinadas a facilitar a recuperação da informação” (CUNHA, 1987, p. 38).

A expressão análise da informação, de origem inglesa, por sua vez, tem nascimento na Biblioteconomia, podendo ser definida, a partir de Lancaster (2004), como um conjunto de procedimentos destinados a expressar o conteúdo de itens bibliográficos.

A análise de assunto (conteúdo) é uma expressão encontrada nas duas áreas que designa as atividades de análise (*stricto sensu*), interpretação e seleção de conteúdos dos documentos, caracterizando-se pela multiplicidade de sentidos e por confusões conceituais. Essa expressão designa, contudo, o conjunto de operações destinadas à extração do conteúdo temático dos documentos (DIAS; NAVES, 2007).

Há, por vezes, uma tendência à compreensão de que a “análise documentária” é mais ampla e que contempla a “análise de assunto” ou a “análise da informação”, uma vez que estas parecem se dedicar tão somente à análise, à interpretação e à seleção de elementos intrínsecos ao documento. Porém, esta compreensão parece não ser a mais adequada para aqueles que adotam essas denominações, principalmente, por considerar que o conteúdo informativo ou a informação compreende o conjunto de estruturas significantes, incluindo os aspectos intrínsecos (temáticos) e extrínsecos (descritivos) dos documentos, justificando o uso corrente da expressão análise da informação.

A produção científica brasileira, na Biblioteconomia e na Documentação, é marcada pela adoção concomitante dos modelos europeu e norte-americano. Na área específica da análise da informação, a literatura não tem se mostrado de forma diferente. A produção científica é caracterizada pela dupla influência e, por conseguinte, as suas construções teórico-conceituais e metodológicas se apresentam de forma híbrida.

Além disso, é forçoso considerar que o estabelecimento de relações interdisciplinares, no campo informacional vem aproximando as temáticas e os conceitos das diferentes áreas que as constituem. Na Biblioteconomia, atualmente, o conceito de informação compreende a noção de **documento**, uma vez que este corresponde ao suporte em que forma e matéria se integram, constituindo o conteúdo informacional. Por outro lado, na Documentação, o conceito de documento não compreende apenas os aspectos extrínsecos, mas também intrínsecos, abarcando o conteúdo informativo. Observa-se, então, que os conceitos e as práticas biblioteconômicas e documentárias evoluíram na mesma direção, em busca de alcançar os conteúdos informativos dos diversos documentos.



Explicativo



Documento

O termo documento deriva do latim **doceo** e **metum**, que designam, respectivamente, ensinar e testemunhar. Assim, o significado denota, ao mesmo tempo, transmissão de conhecimento e instrumento de prova. No século XIX, contudo, *Paul Otlet, ao procurar delimitar o objeto de estudo da Documentação, ampliou o conceito de documento ao defini-lo como uma coisa informativa. O documento passa a ser compreendido como a representação da realidade de forma literária (o livro, a escrita, o texto) e gráfica (o ícone e a imagem). Ademais, há diversas possibilidades de combinações de formas de apresentação do documento, tais como realidade (reália), imagem (ícone), escrita (livro) e representação de uma representação da realidade (desenho, fotografia, gravura e fotografia). Essas possibilidades de combinações dão condições à compreensão do conceito de documento no contexto das tecnologias digitais caracterizado pela presença da multimídia.*

Fonte: OTLET, P. **Traité de Documentation**. Le livre sur le livre. Théorie et pratique. Bruxelles: Éditeurs-Imprimeurs D. van Keerberghen & Fils, 1934.

Os estudos desenvolvidos no campo da Ciência da Informação, nos últimos anos, também têm contribuído para a aproximação da Biblioteconomia, Documentação, Recuperação da Informação e de outras áreas do conhecimento que se dedicam aos diferentes processos informacionais, dentre os quais se destaca a representação da informação.

Nesta unidade, você observará que a análise da informação é abordada como temática e atividade da subárea **tratamento e organização da informação**, que tem como finalidade estudar o conjunto de fundamentos teórico-conceituais e metodológicos necessários à representação, temática e descritiva, dos documentos. A análise da informação é uma atividade basilar que tem por finalidade promover o acesso à informação. A rigor, analisa-se para selecionar, analisa-se para representar e analisa-se para recuperar.

1.2 REPRESENTAÇÃO DA INFORMAÇÃO

A análise da informação é realizada na execução de diferentes atividades de tratamento da informação, tais como catalogação, classificação, indexação e elaboração de resumos. Essas atividades são imprescindíveis à organização e à recuperação da informação porque resultam em diferentes formas de representação e se constituem em pontos de acesso aos documentos.

A representação da informação pode ser compreendida a partir das perspectivas material e formal. Do ponto de vista material, ela corresponde à apresentação concisa do conteúdo do documento. Do ponto de vista formal, ela se refere às atividades de análise, interpretação, seleção e síntese do conteúdo informacional.

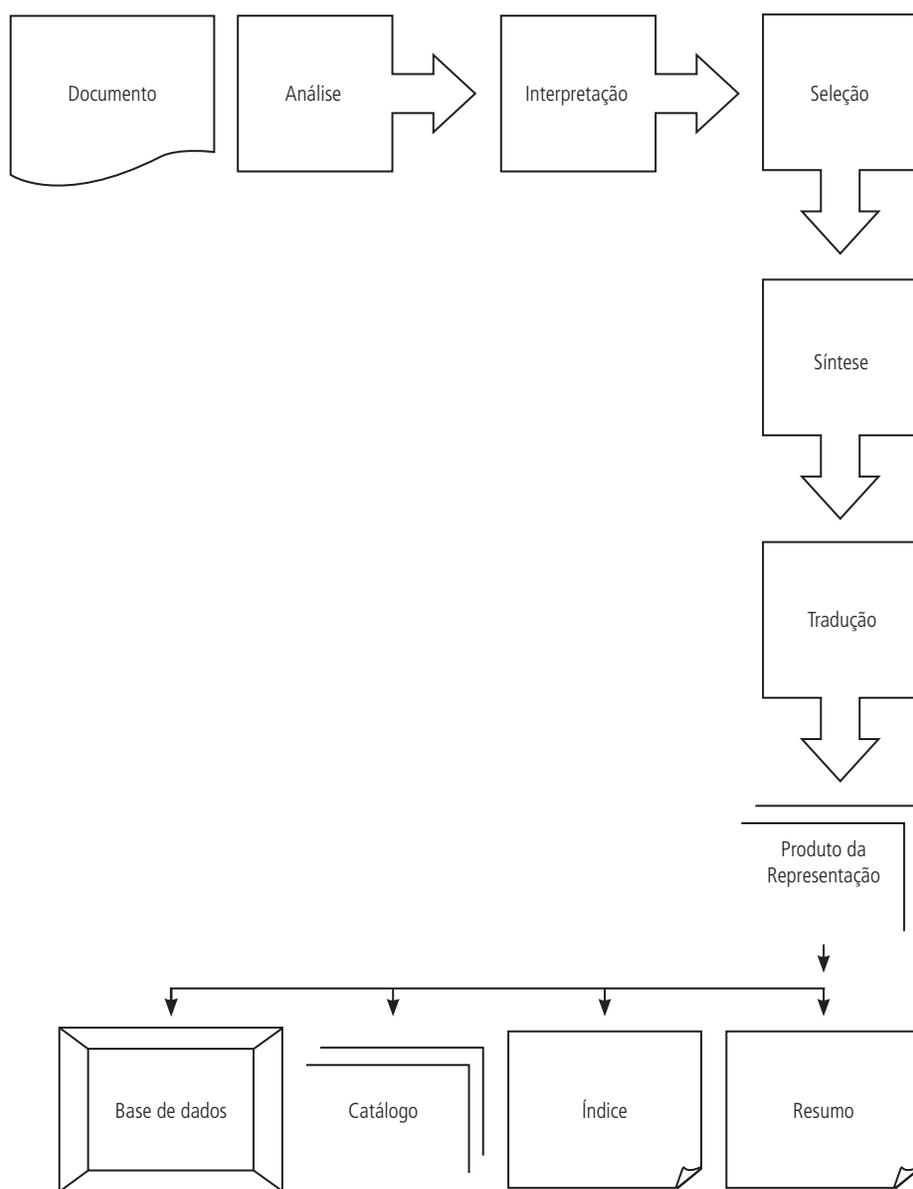
A partir dessa categorização, pode-se perceber que a representação é uma atividade de tratamento da informação que, além de requerer a compreensão da funcionalidade do sistema de informação como um todo, exige outras competências, sobretudo no domínio da leitura e da tomada de decisão. No domínio da leitura, encontram-se as atividades de análise e interpretação dos elementos que melhor representam os documentos, ao passo que, na tomada de decisão, são realizadas as atividades de seleção e síntese dos elementos considerados mais representativos.

O processo de representação da informação se dá, em síntese, a partir de duas etapas:

- a) **análise de conteúdo:** atividade de identificação e definição dos conteúdos informativos dos documentos e
- b) **tradução:** atividade de transformação da linguagem natural adotada nos textos dos documentos na linguagem do sistema, conforme a necessidade de informação do usuário (CESARINO, 1985; LANCASTER, 2004).

A análise da informação é independente do processo de tradução. Existem situações inclusive em que o sistema de informação usa a linguagem natural, não havendo, portanto, necessidade do processo de tradução. Para *Lancaster* (2004, p. 18): "Tradução, a segunda etapa da indexação de assuntos, envolve a conversão da análise conceitual de um documento em um conjunto de termos de indexação".

Figura 1 – Processo de representação da informação



Fonte: Produção do próprio autor (2015).

Como pode ser observado na Figura 1, o processo de representação da informação é realizado a partir de um conjunto de operações consecutivas e resulta em produtos que representam a forma física e/ou o conteúdo dos documentos para posterior identificação, disseminação e/ou recuperação da informação. O tipo de produto e as suas características dependem dos elementos que o compõem e da finalidade da representação. Assim, a representação da informação pode ser classificada, de modo geral, em descritiva e temática.

1.2.1 Representação descritiva da informação

A representação descritiva, também conhecida como representação bibliográfica, corresponde à análise e à síntese de elementos que identificam extrinsecamente o documento, tais como autoria, título, edição, volume, número, local de publicação, editora, ano de publicação, entre outros. Os principais exemplos de representação descritiva da informação são a catalogação e a referência (bibliográfica).

Esse tipo de representação é reconhecido pela presença da objetividade em função do uso de instrumentos auxiliares normativos e, por conseguinte, são também reconhecidos como menos complexos. No processo de catalogação, usa-se, por exemplo, o **Anglo-American Cataloguing Rules** (AACR2) ou segunda edição do Código de Catalogação Anglo-Americano, conforme Figura 2.

Exemplos 1:

Figura 2 – Representação descritiva (ficha catalográfica – AACR2)

020 Le Coadic, Yves-François.
L433c Ciência da Informação / Yves-François Le Coadic; tradução de Maria Yêda Falcão Soares de Filgueiras Gomes. – Brasília, DF : Briquet de Lemos/Livros, 1996.

Título original : La science de l'information.
Inclui bibliografia.
ISBN: 85-85637-08-0

1. Ciência da Informação. I. Título.

Fonte: Produção do próprio autor (2015).

Além do formato em fichas catalográficas, com tamanho padrão de 7,5 x 12,5 cm, a representação descritiva pode ser realizada no formato eletrônico, resultando em representação digital. Nesse caso, o processo de representação tem como referência formatos bibliográficos, como o *Machine Readable Cataloging* (MARC 21), que utiliza protocolos de intercâmbio de dados. Importante destacar que os formatos de intercâmbio contemplam os dados principais da representação descritiva realizada historicamente em formato impresso, conforme pode ser observado na Figura 3.

Exemplo 2:

Figura 3 – Representação descritiva (catálogo eletrônico)

FUNDAÇÃO BIBLIOTECA NACIONAL

Home Pesquisa Autoridades Minha seleção Ajuda

Acessibilidade Alto contraste

Busca rápida Busca combinada Qualquer biblioteca

Autor Euclides da Cunha E - Ano edição a Material Qualquer Idioma Qualquer Ordenação Título - crescente

Buscar Limpar

Últimos títulos acrescentados ao acervo e destaques

igual a

4/208

Detalhes MARC tags Dublin Core

MARC tags

000 00721cam a22002417 4500
001 000218516
003 BR-RJBN
005 20111215133453.9
008 090325:2006 spb 000 1 por
020 _ |a 8572326847 (broch.)
035 _ |a 2009032513483543med
040 _ |a BR-RJBN |b por
082 04 |a 981 |2 22
092 _ |a ANEXO II-811.6.20
100 1 |a Cunha, Euclides da. |d 1866-1909
245 10 |a A margem da história : |b texto integral / |c Euclides da Cunha. -
260 _ |a São Paulo : |b M. Claret, |c 2006.
300 _ |a 234p. : |c 19cm.
490 1 |a (A obra-prima de cada autor : |v 229)
595 _ |a BNB |c 02/09
651 04 |a Brasil |x História
830 0 |a A obra-prima de cada autor |v 229
852 _ |a Obras Gerais
949 _ |a 1.253.030 DL 30/03/2009

Salvar MARC

SophiA

Fonte: Fundação Biblioteca Nacional¹ (2015).

Já na elaboração de referências, são utilizadas a NBR 6023 da Associação Brasileira de Normas Técnicas (ABNT), as *APA's Style Rules* da *American Psychological Association* (APA) e as normas do *International Committee of Medical Journal Editors* (ICMJE), mais conhecido como *Grupo de Vancouver*, entre outras. Embora existam diferenças entre essas normas, sobretudo quanto à formatação, elas contemplam, em grande medida, os mesmos elementos que representam descritivamente o documento. Por outro lado, a NBR 6023/ABNT é direcionada aos diversos tipos de formatos e áreas de conhecimento, ao passo que as normas da APA e de Vancouver são direcionadas aos periódicos científicos, respectivamente, das ciências sociais e comportamentais, e das ciências da saúde. No Brasil, a NBR 6023/ABNT é a mais usada, embora as normas da APA e do ICMJE do Grupo de Vancouver sejam também utilizadas na formatação de comunicações de alguns eventos técnico-científicos e de artigos de alguns periódicos científicos das referidas áreas. A Figura 4 exemplifica a elaboração de uma referência de livro, conforme a NBR 6023/ABNT (2018).

1 FUNDAÇÃO BIBLIOTECA NACIONAL. **Catálogo SOPHIA Web**. Disponível em: <http://acervo.bn.br/sophia_web/>. Acesso em: 03 nov, 2018.

Figura 4 – Representação descritiva (Referência – NBR 6023/ABNT)

LE COADIC, Yves-François. **A Ciência da Informação**. Brasília: Briquet de Lemos Livros, 1996. 119p.

Fonte: Produção do próprio autor (2015).

Resta esclarecer, que, assim como acontece com a catalogação, há diferentes possibilidades de elaboração eletrônica de referências com a adoção de sistemas automatizados. Os detalhes e exemplos desse tipo de recursos eletrônicos para a representação descritiva serão abordados em disciplina que trata especificamente da normalização de trabalhos acadêmicos e técnico-científicos.

1.2.2 Representação temática da informação

A representação temática, também conhecida como intelectual, corresponde à análise e à síntese dos elementos que identificam intrinsecamente os documentos. Trata-se da representação do conteúdo informacional dos documentos. Essa modalidade de representação é subjetiva porque envolve maior capacidade do analista na seleção dos assuntos, principal e secundário(s), na relação com o texto em análise, com o contexto de uso e com a linguagem de indexação, entre outros. Os principais exemplos desse tipo de representação são a classificação, a indexação e a elaboração de resumos, que se encontram amplamente exemplificados na segunda unidade.

Em que pesem as diferenças existentes entre a representação descritiva e a representação temática, esses tipos de representação da informação têm a mesma finalidade, o tratamento, a organização e a recuperação de informações contidas em documentos.



Explicativo

Você sabia?

A Documentação é campo de conhecimento que se dedica ao estudo e à organização dos conteúdos informacionais dos diversos tipos de documentos. Há duas vertentes da Documentação, uma europeia e outra norte-americana. A vertente europeia foi fundada pelo advogado belga *Paul Otlet* (1868-1944) e seus colaboradores, dentre os quais se destaca o também advogado belga *Henri La Fontaine* (1854-1943), em fins do século XIX, a partir da crítica direcionada ao tratamento dado, até então, aos conteúdos dos documentos pela Biblioteconomia. Com base no princípio monográfico, os documentalistas europeus propuseram a análise aprofundada do conteúdo dos documentos e a reorganização das unidades de conhecimento, formando um grande repertório bibliográfico universal, que visava à exclusão dos erros e das duplicações, bem como à aproximação dos conteúdos, conforme as especialidades temáticas. A obra mais conhecida neste campo de conhecimento é o ***Traité de Documentation*** (1934) de autoria de *Paul Otlet*. A vertente norte-americana surgiu também de uma ruptura com a Biblioteconomia e teve por

fundamento a aplicação das tecnologias da microfotografia no tratamento e na recuperação automáticos da informação. Um importante marco foi a criação do **American Documentation Institute** (ADI), em 1932, visando a aglutinar profissionais e instituições preocupados com questões de tratamento e recuperação da informação. A análise da informação, centrada no conteúdo (Europa) e na tecnologia (EUA), correspondeu à essência da evolução dos processos documentais que resultaram na emergência da Documentação.

Fonte: SOUZA, E. D. **A Ciência da Informação: fundamentos epistêmico-discursivos do campo científico e do objeto de estudo.** Maceió, AL: Edufal, 2015. 222 p.

A análise da informação trabalha tanto com os aspectos qualitativos quanto quantitativos. Os primeiros se referem às possibilidades de natureza semântica dos atributos analisáveis e os segundos ao número médio de atributos selecionados e sintetizados. Nesse domínio, definem-se duas medidas de extensão que devem ser consideradas nos processos de análise e síntese da informação, quais sejam a especificidade da linguagem e a exaustividade na atribuição de termos de indexação. Essas medidas têm caráter político e são importantes porque, conforme pode ser observado adiante, interferem diretamente nos coeficientes de recuperabilidade, que, nas palavras de Lancaster (2004), correspondem às medidas de desempenho de um serviço de recuperação da informação.

A especificidade é a medida de extensão usada na análise e síntese da informação que tem como referência a abrangência semântica. O seu aumento pode ser obtido por intermédio de duas operações:

- a) acréscimos de delimitadores ou subcabeçalhos ao assunto genérico. Na Figura 5, por exemplo, "B", "C", "D" e "E" são delimitadores do termo de indexação "A". Cada acréscimo de um desses delimitadores significa um processo de especificação do cabeçalho de assunto.

Exemplos: Biblioteca => Biblioteca **escolar**;
Biblioteconomia = > **História da** Biblioteconomia;
Sistemas de informação = > **Implantação de** sistemas de informação.

- b) substituição do assunto genérico por um assunto específico e/ou acréscimos de cabeçalhos de assuntos mais específicos, conforme o caso. Nesta última possibilidade, contudo, deve, conforme Lancaster (2004), evitar redundâncias.

Exemplos: Arquitetura religiosa => **arquitetura de catedrais**;
Representação da informação => **indexação**;
=> **classificação**;
=> **catalogação**.

O aumento do grau de especificidade com o acréscimo de termos de indexação resulta também no aumento da exaustividade. Veja outro exemplo:

Sistema cardiovascular => **artérias**;
=> **coração**;
=> **vasos capilares**;
=> **veias**.



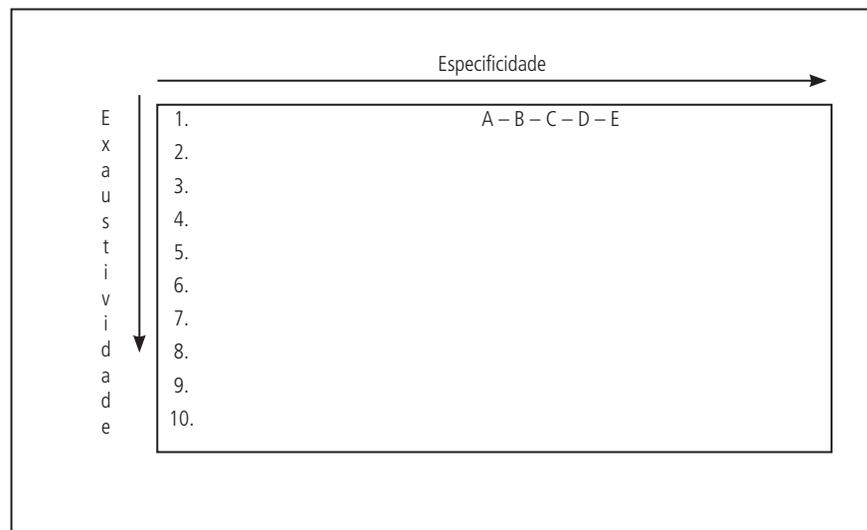
Observe que o processo de especificação se deu pela substituição de um termo genérico (sistema cardiovascular) por quatro termos de entrada específicos (artérias, coração, vasos capilares e veias), resultando, a um só tempo, no aumento tanto do nível de especificidade quanto do nível de exaustividade.

A medida de exaustividade corresponde à abrangência com que o conteúdo de um documento é analisado e representado. Segundo *Lancaster* (2004), a exaustividade refere-se à abrangência da indexação, que, genericamente, corresponde ao número de termos atribuídos ao documento. Assim, quanto mais assuntos forem selecionados no processo de análise e representação da informação, mais exaustiva será a indexação. Da mesma forma, quanto mais termos forem utilizados na indexação de um documento, mais acessível ele se tornará. *Lancaster* (2004, p. 6) esclarece ainda que “[...] os termos atribuídos pelo indexador servem como ponto de acesso mediante os quais um item é localizado e recuperado [...]”. Assim, um documento que trata de diferentes conteúdos pode ser representado por apenas um ou alguns termos de indexação, conforme o nível de exaustividade previamente estabelecido.

- Exemplos: 1º termo de indexação: Biblioteca escolar
- 2º termo de indexação: Biblioteca especial
- 3º termo de indexação: Biblioteca especializada
- 4º termo de indexação: Biblioteca pública
- 5º termo de indexação: Biblioteca universitária

O documento analisado e representado trata dos diversos tipos de bibliotecas. Assim, quanto maior for a quantidade de conceitos selecionados na análise da informação maior o número de termos de indexação e o nível de exaustividade na representação do conteúdo do documento.

Figura 5 – Medidas de extensão da indexação de um documento



Fonte: Adaptado de LANCASTER (2004, p. 30).

É importante ter clareza acerca das medidas de especificidade e exaustividade, que têm como referências dois eixos com orientações distintas, conforme Figura 5. Enquanto a especificidade se encontra no eixo horizontal e define o alcance semântico da linguagem de indexação, a exaustividade está localizada no eixo vertical e se refere à quantidade de termos

atribuídos ao documento. Observe que cada uma dessas medidas tem como opostas, respectivamente, a generalidade e a seletividade. Dessa forma, a representação de um documento pode ser, ao mesmo tempo, específica e exaustiva ou genérica e exaustiva, entre outras combinações.

1.3 ORGANIZAÇÃO DA INFORMAÇÃO

Os sistemas de recuperação da informação, em geral, são planejados e organizados de maneira a promover a eficiente disponibilização e recuperação de informações consideradas relevantes. Para tanto, compreendem um conjunto de atividades consecutivas que vão desde a seleção de documento, para composição de acervos, ao atendimento às necessidades de informação de usuários.

Por algum tempo, as leituras genéricas e superficiais foram suficientes para o tratamento e a organização da informação contida em documentos por duas principais questões: a quantidade de documentos e a finalidade dos acervos documentais. A quantidade se refere tanto à extensão do acervo quanto à diversidade de documentos, e a finalidade refere-se à acumulação. Com a mudança do paradigma de acumulação para o paradigma de acesso, torna-se necessário o desenvolvimento de práticas de tratamento e de organização do conhecimento mais sistematizadas.

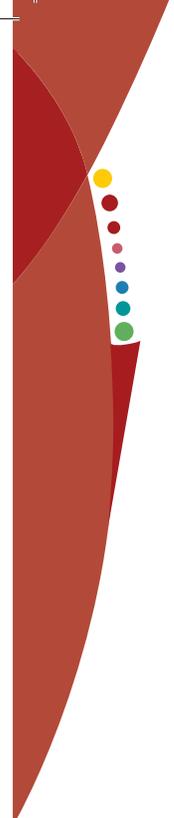
As novas estruturas tecnológicas e informacionais, sobretudo as digitais, ampliaram a capacidade de armazenamento e, essencialmente, os procedimentos de tratamento, organização, disponibilização e recuperação de diversos conteúdos que compõem a crescente massa documental. Esse redimensionamento representou a necessidade de repensar os processos informacionais, como também de construir abordagens teóricas e metodológicas voltadas a diferentes práticas, sobretudo, de elaboração e de aplicação das linguagens documentárias.

Outra questão que tem implicação direta na organização da informação se refere às abordagens adotadas na compreensão dos sistemas de informação que, conforme *Cesarino* (1985), podem ser duas:

- a) **abordagem estruturalista:** considera o sistema de informação como um conjunto de operações consecutivas que visam a promover a recuperação de informações relevantes e
- b) **abordagem sistêmica:** considera o sistema de informação como parte de um modelo de comunicação em um contexto social mais amplo.

Embora, por vezes, pouca ênfase tenha sido dada à segunda abordagem, o fato é que essas abordagens não se excluem, mas se complementam. Não se pode perder de vista que o planejamento e a implementação de sistemas de informação devem considerar o conjunto de atividades que os compõem, a inter-relação entre essas atividades e os contextos institucionais e sociais em que são disponibilizados.





De modo mais preciso, você pode constatar que, conforme *Cesarino* (1985), todos os sistemas de informação são planejados e organizados em três grandes subsistemas:

- a) **entrada:** seleção e aquisição de documentos;
- b) **processamento:** tratamento e organização dos documentos;
- c) **saída:** disseminação (sistema) e recuperação da informação (usuário).

A implantação de um sistema de informação se inicia pelo processo de escolhas de documentos, que tem como base a política de seleção. De acordo com *Vergueiro* (2010), essa política se refere a um conjunto de diretrizes usado no processo decisório, sendo composta por cinco elementos, a saber, critérios de seleção, instrumentos auxiliares, responsável pela seleção, documentos correlatos e políticas específicas. Essas questões serão abordadas em disciplina que trata especificamente da formação e do desenvolvimento de coleções.

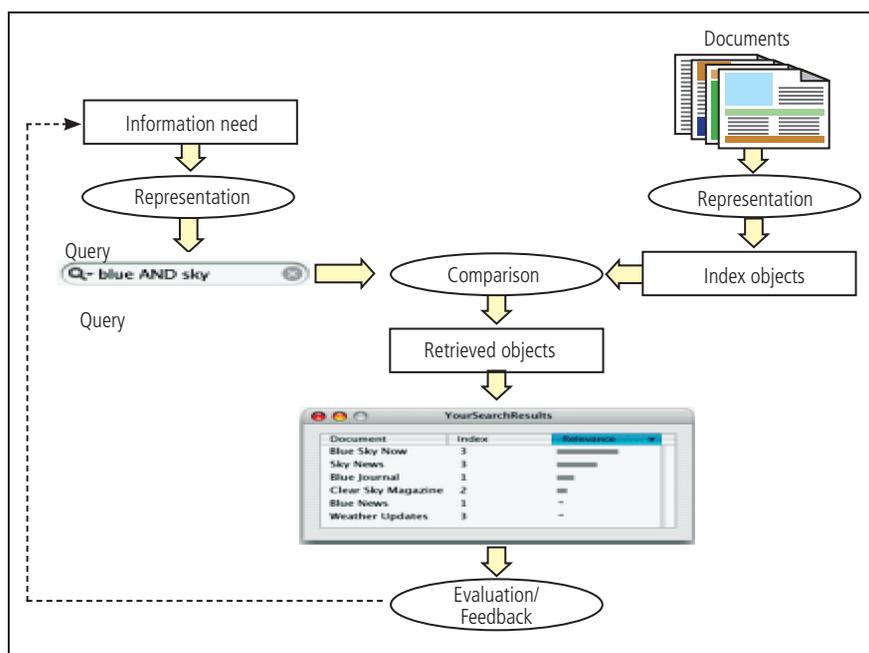
No subsistema de processamento da informação, há um destaque para a política de indexação que corresponde a um conjunto de diretrizes que auxiliam os gestores nas decisões sobre os processos de análise e representação do conteúdo dos documentos. Essas diretrizes abordam questões críticas, tais como os níveis de exaustividade e especificidade. A política é indispensável, sobretudo, em função da impossibilidade de estabelecer regras ou padrões universais e/ou rigorosos para os processos de análise e representação dos conteúdos dos documentos.

A política decide não só sobre a consistência dos procedimentos de indexação em relação aos efeitos que se necessita obter na recuperação, mas, principalmente, sobre a delimitação de cobertura temática em níveis qualitativos e quantitativos tendo em vista os domínios de assuntos e as demandas dos usuários (GIL LEIVA; FUJITA, 2012, p. 17).

Observa-se, conforme *Gil Leiva e Fujita* (2012), que a política de indexação situa as atividades de análise e síntese nos contextos gerencial e estratégico, na medida em que procura planejar um conjunto de princípios e critérios que, a um só tempo, condicionam a adequada representação e a eficiente recuperação da informação.

Nesse contexto, a análise da informação corresponde a um dos principais procedimentos realizados na organização e disponibilização de acervos documentais visando à posterior recuperação de estruturas de informações que os constituem. Essa atividade possibilita uma organização racional e sistemática do conhecimento registrado. Ao mesmo tempo em que esse processo de análise possibilita a organização de acervos, estabelecendo, por exemplo, a localização física ou eletrônica de documentos, condiciona a organização das informações documentais, descritivas e temáticas, em bases de dados e/ou catálogos que serão usado para a localização e o acesso aos documentos.

Figura 6 – Representação da informação no sistema de recuperação da informação



Fonte: MAC DEVELOPER LIBRARY (2005).

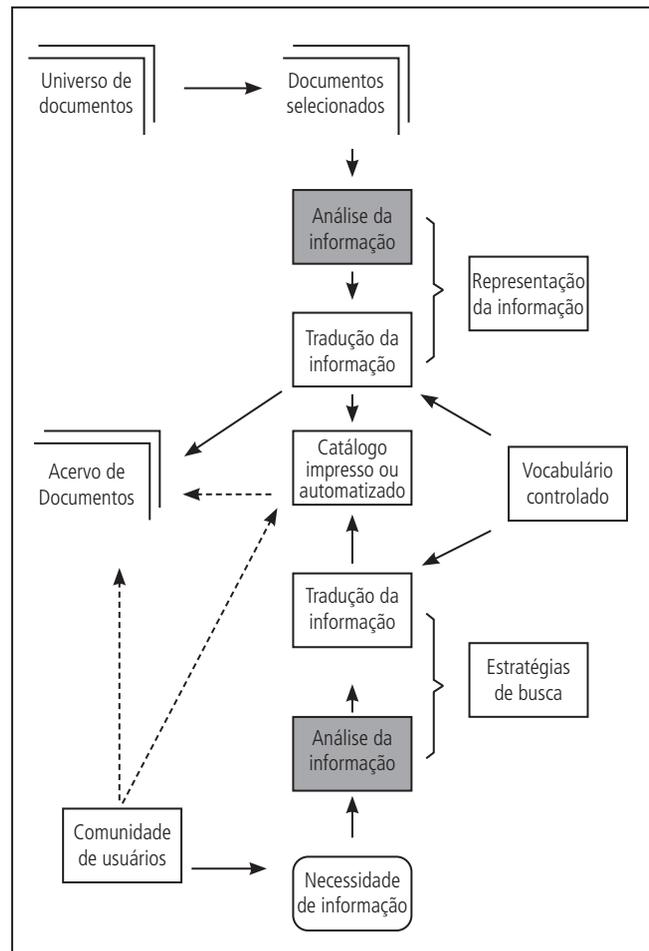
Considerando o modelo sistêmico de tratamento, organização e disseminação/recuperação da informação, observa-se que as atividades de representação e, por conseguinte, de análise e síntese da informação são realizadas em dois principais momentos distintos: representação da informação (indexação e catalogação) e elaboração de estratégias de busca, conforme Figuras 6 e 7. A primeira é realizada pelo bibliotecário, no processamento da informação, e a segunda pelo usuário ou bibliotecário, no momento da recuperação da informação.

A organização da informação tem como finalidade possibilitar o acesso eficiente aos conteúdos informacionais relevantes presentes nos diversos documentos que foram planejadamente selecionados e adquiridos para compor o sistema de recuperação da informação. Esse sistema pode se apresentar de forma também diversa, tais como arquivo, biblioteca, centro de informação, museu, rede virtual de informação, entre outras.

O campo da organização da informação compreende o conjunto de processos, tais como catalogação, classificação, indexação e elaboração de resumos, realizados com base na análise da informação contida em documentos com a finalidade de organização de arquivos dos sistemas de recuperação da informação e de produtos que possibilitam o acesso à informação contida nesses documentos, representados pelos catálogos, índices e resumos, impressos e automatizados, entre outros. Sendo assim, a organização da informação compreende os arranjos físicos e virtuais.



Figura 7 – Análise da informação no sistema de recuperação da informação



Fonte: Adaptado de CESARINO (1985, p. 160) e de LANCASTER (2004, p. 2).

Do ponto de vista do usuário, o efetivo acesso à informação implica identificação, localização e obtenção de documentos diversos. Nesse universo, o uso das linguagens documentárias se apresenta como fundamental. As redes de relações semânticas que são estabelecidas entre os conceitos, nas linguagens documentárias, são fundamentais para a representação e recuperação da informação, como também para que a disseminação do conteúdo informacional de documentos esteja além da sua unidade material (conteúdo) e formal (arranjo).

1.4 RECUPERAÇÃO DA INFORMAÇÃO

A análise da informação é uma atividade essencial a todo sistema de recuperação da informação porque ela estabelece os ajustes necessários à comunicação entre o usuário e o sistema. Sem esses ajustes, o processo de recuperação da informação torna-se impossibilitado ou, pelo menos, prejudicado. Isso quer dizer que a qualidade do sistema de recu-

peração da informação depende muito da forma como usuário e sistema interagem e, como nos demais processos de comunicação, da linguagem utilizada.

Se não houver um ajuste terminológico entre o universo do sistema e o universo do usuário, o processo de comunicação torna-se impraticável. Esse ajuste se dá, na maioria das vezes, por intermédio do controle linguístico e da definição de uma rede semântica entre termos e conceitos que sistematizados integram a terminologia especializada do domínio em análise.

A análise da informação contida em documentos e nas questões de busca de usuários resulta em um conjunto de elementos representativos que constituem os pontos de acesso para recuperação da informação. Esse conjunto de elementos preferenciais corresponde à linguagem utilizada pelo sistema.

Como abordado anteriormente, no domínio do subsistema de processamento, a análise da informação é realizada com o objetivo de representar os documentos e com a finalidade de recuperá-los posteriormente. A representação da informação pode ser realizada de duas formas, a saber, temática e descritiva. Essas duas formas de representação possibilitam, a um só tempo, a organização dos diversos documentos que compõem os acervos dos sistemas e a recuperação das informações contidas nos documentos.

Conforme visto acima, o processo de representação da informação tem na sua base as duas medidas de extensão, a especificidade e a exaustividade. Essas medidas mantêm relação direta com a recuperação da informação, uma vez que condicionam os índices de eficiência do sistema de informação, que são denominados de coeficientes de recuperabilidade.

De modo geral, o coeficiente de recuperabilidade tem como finalidade avaliar o nível de desempenho do sistema de recuperação da informação, considerando o conteúdo dos documentos e as necessidades de informação dos usuários, e é classificado em coeficiente de precisão e coeficiente de revocação (LANCASTER, 2004; PIEDADE, 1983).

O coeficiente de precisão corresponde à capacidade que o sistema de recuperação de informação tem de recuperar documentos relevantes. Em outros termos, *Lancaster* (2004) esclarece que o seu cálculo é realizado a partir da relação direta entre os documentos relevantes recuperados e o total de documentos recuperados pelo sistema, conforme fórmula apresentada na Figura 8.

Figura 8 – Coeficiente de precisão

$$\text{Coeficiente de precisão} = \frac{\text{Documentos relevantes recuperados}}{\text{Total de documentos recuperados}}$$

Fonte: Produção do próprio autor (2015).

Assim, em uma situação prática, no processo de busca sobre a temática “índice pluviométrico” realizado em um sistema de recuperação de informação são recuperados 150 (cento e cinquenta) documentos, sendo que destes apenas 15 (quinze) se apresentam como relevantes para o usuário: o índice de precisão é de 0,10, ou seja, 10%.



Observe que esse índice de recuperabilidade não tem como centralidade a preocupação com a totalidade de documentos relevantes sabidamente existentes no sistema de informação, mas busca evitar a recuperação de documentos que não atendam às necessidades de informação do usuário. Em outros termos, em um sistema de informação que tenha como meta de desempenho a precisão, tem-se como preocupação evitar os ruídos no processo de recuperação de informação.

A partir dessa perspectiva, procura-se correlacionar a medida de extensão da especificidade com o coeficiente de recuperabilidade da precisão. O fato é que existe uma relação direta entre essas duas medidas, ou seja, quanto maior for o grau de especificidade da linguagem adotado na representação da informação, maior será o índice de precisão na recuperação de documentos. A especificidade, portanto, auxilia na diminuição dos ruídos no processo de recuperação da informação.

O coeficiente de revocação, por sua vez, corresponde à extensão com que todos os documentos relevantes sabidamente existentes no sistema de informação são recuperados. O seu cálculo é realizado a partir da relação direta entre os documentos relevantes recuperados e os documentos relevantes disponíveis no sistema (LANCASTER, 2004; PIEDADE, 1983), conforme fórmula apresentada na Figura 9.

Figura 9 – Coeficiente de revocação

$$\text{Coeficiente de revocação} = \frac{\text{Documentos relevantes recuperados}}{\text{Total de documentos relevantes existentes no acervo}}$$

Fonte: Produção do próprio autor (2015).

Na mesma situação prática acima, considerando que foram recuperados 15 (quinze) documentos relevantes para o usuário, no universo de 30 (trinta) documentos relevantes existentes no sistema de informação, o índice de revocação foi de 0,50, ou seja, 50%.

Observe que, em um sistema de informação que adota esse coeficiente de recuperabilidade, tem-se como centralidade a preocupação em recuperar a totalidade de documentos relevantes disponíveis, não se preocupando com a possibilidade de recuperar também documentos que não atendem às necessidades do usuário. Em outros termos, em um sistema de informação que tem como meta o alto índice de revocação, procura-se evitar o silenciamento, isto é a ausência de resposta do sistema, no processo de recuperação da informação.

É também em função desse entendimento que se procura correlacionar a medida de extensão da exaustividade com o coeficiente de recuperabilidade da revocação. Com efeito, há uma relação direta entre essas duas medidas, ou seja, quanto maior for o nível de exaustividade adotada na análise e na representação da informação maior será o índice de revocação na recuperação dos documentos.

O aumento do nível de exaustividade, acompanhado do aumento do nível de especificidade, resulta também no aumento da precisão. Essa situação ocorre quando a exaustividade é obtida pelo processo de substituição de um termo genérico por um conjunto de termos mais específicos. A exaustividade, portanto, auxilia na diminuição dos silêncios no processo de recuperação da informação.

Há alguns elementos que condicionam a eficiência do processo de análise da informação:

- a) disponibilidade de recursos humanos;
- b) disponibilidade de recursos tecnológicos;
- c) tipos de arquivos usados pelo sistema de informação;
- d) linguagens utilizadas – coordenadas ou pós-coordenadas e
- e) nível de acesso ao usuário (CESARINO, 1985).

A definição política dos coeficientes de recuperabilidade tem por base o estabelecimento da relevância dos conteúdos dos documentos para a comunidade de usuários a que se destinam. Esta se torna uma atividade bastante complexa porque envolve diferentes elementos que interferem na compreensão e na definição dos conteúdos dos documentos e das necessidades de informação dos usuários.

RESUMO

A análise da informação é abordada como temática e atividade da **subárea tratamento e organização da informação**, orientada aos processos de busca e recuperação da informação, em bibliotecas, centros de documentação e centros de informação, entre outros. Os estudos da análise da informação são caracterizados pelo pluralismo teórico-conceitual em decorrência da dupla influência: Documentação (europeia) e Biblioteconomia (norte-americana). A **representação da informação** é uma atividade de tratamento realizada a partir de duas etapas, **análise de conteúdo** (análise documentária ou análise da informação) e **tradução**.

Esse processo é realizado a partir de um conjunto de operações consecutivas e resulta em produtos que representam a forma física (**representação descritiva**) e o conteúdo do documento (**representação temática**), que tem como finalidade a busca e a recuperação da informação. A análise da informação compreende os aspectos extrínsecos e intrínsecos do documento, pois possibilita a identificação e a seleção dos conteúdos informativos dos documentos, das linguagens documentárias e do campo semântico em que estes se situam. Na representação temática, essa atividade abarca aspectos qualitativos e quantitativos, definindo as medidas de extensão, a saber, **especificidade** da linguagem (alcance semântico de termos) e **exaustividade** da representação (quantidade de termos). As atividades de tratamento, descritivo e temático, condicionam a **organização da informação** nos diferentes sistemas de recuperação da informação porque têm como objetivo a organização racional e sistemática do conhecimento registrado, bem como das informações documentárias usadas na localização dos conteúdos informativos.

A organização da informação, de modo geral, compreende os processos de catalogação, classificação, indexação e elaboração de resumos, entre outros, que resultam na organização dos documentos e das informações documentárias. Estas, por sua vez, são estruturadas em produtos documentários, como base de dados, catálogos, índices e resumos, que

Relevância – A eficiência dos sistemas de informação tem como referência os coeficientes de recuperabilidade de documentos considerados relevantes pelos usuários. A relevância, nesse sentido, é essencialmente o termômetro dessa eficiência, que corresponde, em tese, à correlação entre os conteúdos informativos dos diversos documentos e as demandas dos usuários. Trata-se, portanto, de um julgamento do usuário que, como tal, está condicionado a uma série de fenômenos que abarcam inclusive questões complexas, tais como emocionais e afetivas. A relevância se apresenta como uma das principais categorias de estudos atuais sobre a eficiência dos sistemas de recuperação da informação.

Fonte: LANCASTER, F. W. **Indexação e resumo**: teoria e prática. 2. ed. Brasília: Briquet de Lemos Livros, 2004. 452 p.

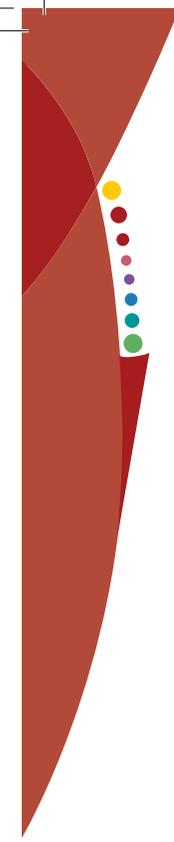


V. a representação temática tem como base os elementos intrínsecos do documento e se caracteriza pela subjetividade.

Estão corretas:

- a) I e II, apenas.
 - b) I e III, apenas.
 - c) I, II, III e IV.
 - d) III e IV, apenas.
 - e) II e IV, apenas.
3. A representação da informação é uma das atividades principais de uma biblioteca, pois permite a organização e a recuperação da informação. Pode-se afirmar assertivamente que, de modo geral, ela é realizada a partir de duas etapas básicas.
- a) Interpretação e síntese.
 - b) Análise e interpretação.
 - c) Análise e tradução.
 - d) Interpretação e seleção.
 - e) Seleção e síntese.
4. A análise da informação é realizada, pelo menos, em dois momentos distintos no sistema de recuperação da informação. A esse respeito, aponte a alternativa correta.
- a) A adequação dos documentos aos critérios de seleção é a principal atividade analítica adotada na representação da informação.
 - b) A representação da informação e a elaboração de estratégias de buscas são os dois momentos em que a análise da informação e a tradução estão presentes.
 - c) Na elaboração de estratégias de busca, o controle de vocabulário é realizado exclusivamente pelo usuário.
 - d) A interpretação e a síntese são atividades de controle de vocabulários presentes, respectivamente, na representação e nas estratégias de busca.
 - e) Na representação da informação, a análise é condicionada à necessidade do processo de tradução.
5. As medidas de extensão e os coeficientes de recuperabilidade estão diretamente relacionados à eficiência dos sistemas de recuperação da informação. Analise as assertivas e marque a alternativa correta.
- I. A revocação, que é um coeficiente de recuperabilidade, reduz a capacidade de recuperação do sistema de documentos irrelevantes.
 - II. A exaustividade, que é uma medida de extensão, amplia a capacidade do sistema de recuperação de documentos relevantes.
 - III. A precisão, que é uma medida de extensão, aumenta a capacidade do sistema de recuperação de documentos relevantes.





IV. A especificidade, que é uma medida de extensão, é obtida com o acréscimo de subcabeçalhos ou de cabeçalhos mais específicos.

Estão corretas:

- a) I e II, apenas.
 - b) II e III, apenas.
 - c) I e III, apenas.
 - d) I, II, III e IV.
 - e) II e IV, apenas.
-

UNIDADE 2

PROCEDIMENTOS DE ANÁLISE DA INFORMAÇÃO





A análise da informação é um processo de tratamento da informação caracterizado pela complexidade, principalmente, em função da subjetividade que a envolve. Conforme você já observou anteriormente, trata-se de identificar e determinar o conteúdo informativo, temático e descritivo de um documento. Essa definição tem na sua base um conjunto de operações intelectuais que resultam na interpretação e na seleção daqueles conteúdos considerados representativos do documento analisado.

A representação temática da informação pode ser realizada fazendo uso da linguagem natural ou de uma linguagem artificial específica. O uso dessas linguagens apresenta particularidades que estão diretamente relacionadas à dispersão terminológica e à dispersão sintática, próprias da linguagem natural. Assim, ingressar no universo do controle terminológico e sintático da linguagem natural significa necessariamente estabelecer um conjunto de relações entre os elementos que compõem a estrutura representacional.

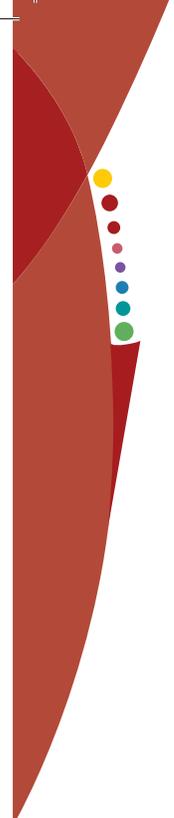
Para melhor compreensão dos procedimentos específicos adotados na análise da informação, de forma particular, e na representação, de modo geral, é imprescindível você compreender os elementos que compõem essas linguagens, os instrumentos usados no processo de representação e os produtos utilizados na recuperação da informação. Nesse horizonte, esta unidade tem como objetivo apresentar as condições e o conjunto de procedimentos que compõem os processos de análise e síntese da informação, que envolve linguagens, instrumentos e produtos usados no tratamento e na organização da informação.



2.1 LINGUAGEM NATURAL E LINGUAGEM ARTIFICIAL

Como foi dito anteriormente, a linguagem é a base da comunicação entre usuário e sistema de informação. Há dois domínios de informação que precisam estar acessíveis aos usuários, um mediato e outro imediato. O primeiro corresponde ao conjunto de conteúdos informativos que compõem os diversos documentos e o segundo, às informações sobre esses documentos que são organizados e disponibilizados pelo sistema de informação. Estas últimas são denominadas "informações documentárias", "informações referenciais" (descritivas e temáticas), "metainformação" etc. Têm-se aí dois universos distintos que precisam ser aproximados para que se obtenham resultados satisfatórios na esfera comunicacional e, finalmente, no efetivo uso do sistema de recuperação da informação: o universo dos conteúdos registrados nos documentos e o universo das necessidades de informação expressos nas demandas dos usuários. Você pode, assim, perceber que o sistema de recuperação da informação tem um papel fundamental de mediação entre esses dois universos.

Por um lado, os conteúdos informacionais são as diversas expressões de seus respectivos autores e, por outro, as demandas informacionais são



diferentes manifestações de necessidades informacionais dos usuários. O sistema de recuperação da informação procura, portanto, estabelecer a efetiva comunicação entre dois universos povoados de singularidades, notadamente, no modo de expressar os conteúdos informativos que formam os documentos e as necessidades de informação. Essa diversidade se manifesta por intermédio das dispersões semânticas e sintáticas, próprias da linguagem natural.

Assim, essa dinâmica é a essência do processo comunicacional entre seres humanos, que são dotados de subjetividades. Em função dessa subjetividade e das duas modalidades de dispersão, um mesmo conteúdo informacional é expresso em diferentes documentos de formas distintas.

O sistema de recuperação da informação precisa, portanto, estabelecer as condições mínimas para que os atores do processo comunicacional dialoguem, em um ambiente de mediação, que, em tese, se apresenta de forma bastante estruturada para atender às necessidades de informação do usuário. Sob essas perspectivas, no domínio da representação da informação, a linguagem é classificada como natural ou artificial.

A **linguagem natural**, como a própria expressão designa, corresponde ao conjunto de significantes, organizados semântica e sintaticamente, usados pelos seres humanos no processo comunicacional. De modo mais preciso, corresponde à linguagem adotada pelos autores na composição dos conteúdos informacionais e pelos usuários na elaboração de suas demandas de informação.

Considera-se a expressão **linguagem natural** como sinônimo de “discurso comum”, isto é, a linguagem utilizada habitualmente na escrita e na fala, e que é o contrário de “vocabulário controlado”. No contexto da recuperação da informação, a expressão normalmente se refere às palavras que ocorrem em textos impressos e, por isso, considera-se como seu sinônimo “texto livre” (LANCASTER, 2004, p. 250, *grifos do autor*).

O processo de representação da informação usando a linguagem natural é denominado de **indexação por extração**, **indexação derivativa** ou **indexação por palavras**, uma vez que são usadas as palavras ou expressões adotadas pelo autor no título ou no texto. O principal exemplo de indexação por palavras é o índice *Keyword in context* (KWIC), que utiliza o título dos documentos no processo de indexação.

O índice KWIC [...] é um índice rotado, derivado, em sua forma mais comum, dos títulos de publicações. Cada **palavra-chave** que aparece num título torna-se ponto de entrada, destacada de alguma forma, aparecendo, normalmente, realçada no centro da página [...] (LANCASTER, 2004, p. 54, *grifo do autor*).

Da mesma forma que a linguagem natural, as linguagens documentárias se constituem em sistemas simbólicos que têm como objetivo permitir o processo de comunicação entre, pelo menos, dois universos, a saber, do emissor (mundo do autor e do sistema/editorial) e do receptor (mundo do usuário da informação). A diferença se encontra, contudo, nos uni-

versos de referência; enquanto a linguagem natural tem como referência o mundo social, como um todo, a linguagem documentária tem como referência o mundo documentário ou, em outros termos, o sistema de recuperação da informação, especificamente, a relação usuário-sistema (CUNHA, 1987; DODEBEI, 2002).

As **linguagens documentárias** correspondem a um conjunto de termos e/ou símbolos destinados à representação dos conteúdos dos documentos em sistemas de recuperação de informação. Cada linguagem documentária se organiza em função de suas finalidades e dos procedimentos adotados nos processos de análise, representação e recuperação da informação. Conforme *Piedade* (1983), a linguagem empregada na representação dos conteúdos dos documentos é também denominada de **linguagem descritora, linguagem de indexação e linguagem de informação**, refletindo, em certa medida, a dispersão terminológica já abordada na unidade anterior.

Você observará que, como as demais, as linguagens documentárias são compostas por dois componentes básicos, a saber, o **vocabulário** e a **sintaxe**. Essa estrutura básica das linguagens documentárias possibilita a dinâmica do processo de representação da informação, principalmente por considerar que a linguagem natural evidencia em sua essência a subjetividade. Com efeito, as linguagens documentárias são planejadas e elaboradas para o uso exclusivo no universo documentário. Por isso, cada significante corresponde a um significado ou a um conjunto de significado previamente estabelecido por intermédio do controle de vocabulário.

Assim como acontece com o processo analítico, as linguagens documentárias atuam nos sistemas de recuperação da informação em dois momentos, na representação e na recuperação da informação. No processo de representação da informação, a linguagem documentária orienta o bibliotecário sobre os elementos mais indicados para representar os conteúdos dos documentos. No processo de recuperação da informação, ela orienta o usuário, em estado de busca, sobre quais os elementos que melhor representariam o conteúdo do documento que é objeto de pesquisa. Em outros termos, a linguagem documentária orienta o usuário na elaboração de suas estratégias de busca.

As linguagens documentárias têm como objetivo fazer o controle de vocabulário da linguagem natural, permitindo um melhor diálogo entre o sistema de informação e o usuário; em última análise, a recuperação da informação. O controle de vocabulário se dá tanto na esfera semântica quanto na sintática. O processo de representação realizado com base em linguagens documentárias é denominado de **indexação por atribuição** ou **indexação por conceitos**.

No domínio semântico, esse controle é realizado a partir do estabelecimento de diferentes relações entre os termos que compõem a linguagem. As relações semânticas, conforme *Pinto* (1985), podem ser classificadas em três categorias: **relações de equivalência, relações hierárquicas e relações afinitivas** ou **associativas**.

A definição de relações de equivalência tem como objetivo estabelecer termos substitutivos usados na representação da informação. Existem várias subcategorias desse tipo de relação, dentre as quais se destacam:



a) **sinônimos:** termos diferentes que possuem o mesmo significado.

Exemplos:

disfunção x distúrbio	retrato x fotografia	tangerina x bergamota
-----------------------	----------------------	-----------------------

b) **antônimos:** termos que apresentam significados contrários.

Exemplos:

alegria x tristeza	conformismo x resiliência	introspectivo x comunicativo
--------------------------	---------------------------------	------------------------------------

c) **grafias diferentes:** termos grafados ligeiramente de forma diversa, mas com o mesmo significado.

Exemplos:

contacto x contato	loira x loura	secção x seção
--------------------	---------------	----------------

d) **abreviaturas e siglas:** representação de termos ou expressão de forma reduzida.

Exemplos:

IBICT x Instituto Brasileiro de Informação em Ciência e Tecnologia	km x quilômetro	MEC x Ministério da Educação
--	-----------------------	------------------------------------

e) **traduções:** termos ou expressões grafados em idiomas diferentes.

Exemplos:

Information Retrieval x Recuperação da Informação	Information Science x Ciência da Informação	recall x revocação
---	--	---------------------------------

As relações de equivalência são restritivas, uma vez que especificam os termos preferenciais que foram usados na representação e que, por conseguinte, devem ser usados na busca da informação. Essas relações são do tipo **ver**, representada por "x", ou **use/usado para**.

A definição de relações hierárquicas possibilita a busca por termos alternativos usados na representação da informação. Esse tipo de relação amplia a possibilidade de estratégias de pesquisa no sistema de recuperação da informação. Existem duas subcategorias desse tipo de relação:

a) **gênero e espécie:** designa a relação de grupo de elementos que têm a mesma aparência (do latim *species*) e elementos deste grupo.

Exemplos:

felino xx leopardo	vertebrado xx cavalo	veículo xx carro
--------------------------	----------------------------	------------------------



- b) **todo e parte:** expressa a relação entre a unidade de um elemento e uma fração deste.

Exemplos:

aparelho circulatório xx veias	escola xx sala de aula	planta xx raiz
--------------------------------------	------------------------------	----------------------

As relações lógicas ou hierárquicas de gênero e espécie são ampliativas, uma vez que definem outros termos, mais específicos e/ou mais genéricos, que podem ser adotados na representação e que, por conseguintes, podem ser usados na busca da informação. Essas relações ampliam as estratégias de busca e são do tipo **ver também**, representada por “xx” (CAMPOS, 2001; PINTO, 1985).

A definição de relações afinitivas ou associativa também possibilita a busca por termos alternativos usados na representação da informação. Assim, esse tipo de relação também amplia as possibilidades de estratégias de pesquisa no sistema de recuperação da informação. Há, conforme *Pinto* (1985), várias possibilidades de relações afinitivas, dentre as quais se destacam:

- a) **causa e efeito:** designa a causalidade entre dois fenômenos consecutivos, sendo que o segundo é consequência do primeiro.

Exemplos:

ensino xx aprendizagem	exercício físico xx hipertrofia muscular	vacina xx imunidade
------------------------------	--	---------------------------

- b) **concorrência:** designa a relação entre dois elementos que convergem para um mesmo ponto.

Exemplos:

educação xx ensino	procura xx oferta	trabalho xx produção
--------------------------	-------------------------	----------------------------

- c) **coordenação:** estabelece relação entre dois elementos independentes, mas que participam do mesmo domínio sistêmico ou hierárquico.

Exemplos:

arquivo xx biblioteca	porta xx janela	reino animal xx reino vegetal
-----------------------------	-----------------------	-------------------------------------

- d) **genética:** condiciona a relação entre elementos em decorrência de determinações genéticas.

Exemplos:

infarto xx gengivite	mastite clínica xx células somáticas	pai xx filho
----------------------------	--	--------------------



e) **instrumental:** expressa a relação entre instrumentos e processos.

Exemplos:

AACR2 xx catalogação	Tabela de Cutter xx notação de autor	tesauro xx indexação
----------------------------	--	----------------------------

f) **similaridade de processos e produtos:** designa processos e produtos que são realizados adotando um conjunto de procedimentos semelhantes.

Exemplos:

classificação xx indexação	índice xx catálogo	índice xx resumo
----------------------------------	--------------------------	------------------------

g) **substância:** evidencia a relação entre o produto e a substância de que é constituído.

Exemplos:

bomba atômica xx urânio	etanol xx cana de açúcar	papel xx madeira
-------------------------------	--------------------------------	------------------------

As relações afinitivas também são ampliativas, na medida em que definem outros termos afins que podem ser usados na representação e, por conseguinte, na busca de informação. Essas relações similarmente ampliam as estratégias de busca e são do tipo **ver também**.

Nesse domínio, é importante também esclarecer a diferença entre relações hierárquicas e relações lógicas e ontológicas. Dessa forma, conforme *Campos* (2001), as relações lógicas e ontológicas têm como finalidade estabelecer a natureza da relação entre conceitos, sem, no entanto, definir ordem entre eles. As relações hierárquicas, por sua vez, além de estabelecer a relação entre conceitos, têm por finalidade definir uma ordem sistemática entre os conceitos, que resultam na formação de uma estrutura classificatória.

Campos, Gomes e Motta (2004) defendem que existe uma relação de subordinação lógica quando a intensão do conceito subordinado inclui a intensão do conceito superordenado e, pelo menos, uma característica especificadora. Por outro lado, as relações ontológicas incluem as relações partitivas e associativas.

Embora não haja uma uniformidade, na literatura, sobre essa classificação, *Campos* (2001) esclarece que, na Teoria da Classificação e na Teoria da Terminologia, as relações do tipo gênero-espécie são consideradas relações lógicas, na medida em que são estruturadas a partir do agrupamento de renques e cadeias, que possibilitam a organização da estrutura classificatória. Isso significa que, além de estabelecer relações entre conceitos, como ocorrem com as relações lógicas e ontológicas, as relações hierárquicas têm como finalidade posicionar os conceitos em uma estrutura sistemática vertical.

Os renques se referem a classes derivadas de um universo com referência em apenas uma característica, visando ao estabelecimento de um arranjo e à formação de séries horizontais. A cadeia, por sua vez, concerne a uma sequência formada por classes e seu universo de deslocamento, organizadas de forma vertical. Cada classe que compõe a cadeia tem uma característica a mais ou a menos, conforme a escala, ascendente ou descendente (CAMPOS, 2001; RANGANATHAN, 1967).

Exemplo de renques:

Reino animal

Reino vegetal

Exemplo de cadeia:

Animalia (**reino**)

Arthropoda (**filó**)

Chelicerata (**subfiló**)

Arachnida (**classe**)

Scorpiones (**ordem**)

Euscorpiidae (**família**)

Megacormus (**gênero**)

Megacormus granosus (**espécie**)

Considerando o conjunto de relações possíveis entre palavras ou termos que representam o conteúdo de um documento, destacam-se as relações de coordenação, que são bastante presentes nas diversas linguagens documentárias. Essas linguagens podem ser classificadas em pré-coordenadas e pós-coordenadas, em relação à representação, segundo o momento em que os termos ou símbolos são coordenados para formar o cabeçalho de assunto ou o descritor.

As **linguagens pré-coordenadas** são aquelas que coordenam ou combinam os termos ou símbolos que representam os conteúdos dos documentos no momento da representação da informação. Assim, o processo de coordenação que se dá na conformação de assuntos compostos e complexos é realizado pelo bibliotecário no momento da representação da informação. É importante esclarecer que, conforme Campos (2001), quando o assunto de um documento é formado pela combinação de um assunto básico com um ou mais isolados, é denominado assunto composto. E, quando aquele é formado pela combinação de dois assuntos básicos, denomina-se assunto complexo.

Exemplo de assunto composto:

Introdução à Biblioteconomia no Brasil.

Biblioteconomia (assunto básico)

Brasil (assunto isolado)

Introdução (assunto isolado)

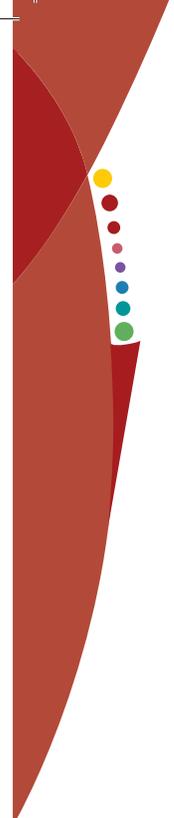
Exemplo de assunto complexo:

Estatística aplicada à Biblioteconomia

Biblioteconomia (assunto básico)

Estatística (assunto básico)





A pré-coordenação na representação da informação tem como objetivos:

- a) a formação de conceitos compostos e complexos;
- b) a predeterminação de relações entre conceitos;
- c) a definição da ordem de citação;
- d) a especificação do conteúdo;
- e) a exclusão de falsas coordenações e
- f) a melhor precisão no processo de recuperação da informação.

A formação de conceitos compostos e complexos resulta da combinação de, pelo menos, dois termos que expressam conceitos simples. Essa combinação é realizada no momento da análise e compreende a predeterminação de relações entre conceitos, considerando a integração sistema-usuário.

O estabelecimento da ordem de citação é a definição da posição dos assuntos na composição de cabeçalhos compostos e complexos. Trata-se de uma tomada de decisão que hierarquiza a ordem de prioridade em que os conceitos aparecem na conformação do cabeçalho de assunto, definindo, assim, a função de cabeçalho e subcabeçalho para cada assunto.

A especificidade do conteúdo de um documento, conforme já visto na primeira unidade, é realizada pelo acréscimo de termos delimitadores, quase sempre denominados números auxiliares ou subcabeçalhos, respectivamente, nos sistemas de classificação e nos cabeçalhos de assunto.

A exclusão de falsas coordenações está diretamente relacionada à prefixação da ordem de citação, que, ao estabelecer a posição dos assuntos na composição dos cabeçalhos compostos e complexos, retira a autonomia do usuário de realizar outros tipos de construções.

O aumento do grau de precisão no processo de recuperação da informação se dá com a delimitação do alcance semântico da representação, que é realizado pela definição do nível de especificidade da linguagem como medida de extensão.

As linguagens pré-coordenadas são denominadas de linguagens tradicionais porque são comumente usadas em sistemas de recuperação da informação manuais e procuram, na medida do possível, aproximar os documentos às demandas de seus respectivos usuários, por intermédio do controle sistemático (CESARINO; PINTO, 1978).

Os principais exemplos dessas linguagens são os sistemas de classificação bibliográfica e as listas de cabeçalhos de assunto. Os sistemas de classificação mais conhecidos e utilizados no Brasil são a *Classificação Decimal de Dewey* (CDD), a *Classificação Decimal Universal* (CDU) e a *Classificação da Library of Congress* (LC). As listas de cabeçalhos de assunto mais conhecidas e utilizadas no Brasil são a *Library of Congress of Subject Headings* (LCSH) e a *Sears List of Subject Headings* (Sears). Essas linguagens são, portanto, instrumentos auxiliares adotados no processo de representação da informação, conforme serão abordados mais à frente.

A seguir, são apresentados exemplos de assuntos pré-coordenados, usando os diferentes instrumentos auxiliares.

Exemplo 1: Direito previdenciário brasileiro

Na linguagem natural ou no enunciado do texto-fonte, tem-se:

1º conceito: Direito

2º conceito: Previdência

3º conceito: Brasil

Quadro 1 – Tipos de representação temática

Instrumento	Notação
Cabeçalho de assuntos	DIREITO PREVIDENCIÁRIO – BRASIL
CDD	344.030981
CDU	349.3(81)

Fonte: Produção do próprio autor (2015).

Observe, no Quadro 1, que a pré-coordenação dos termos simples (“Direito”, “Previdência” e “Brasil”) promoveu, além da definição da ordem de citação, a exclusão de falsas coordenações, a especificação da representação e a probabilidade de precisão na recuperação da informação. Trata-se de um documento que tem como assunto geral “Direito”, mas que foi especificado pelo acréscimo de dois delimitadores, a saber, “previdenciário” e “Brasil”.

Exemplo 2: História da Biblioteconomia

Na linguagem natural ou no enunciado do texto-fonte, tem-se:

1º conceito: História

2º conceito: Biblioteconomia

Quadro 2 – Tipos de representação temática

Instrumento	Notação
Cabeçalho de assuntos	BIBLIOTECONOMIA – HISTÓRIA
CDD	020.9
CDU	02(091)

Fonte: Produção do próprio autor (2015).

Observe, no Quadro 2, que a coordenação dos termos simples (“História” e “Biblioteconomia”), além de promover a formação do conceito composto, estabeleceu a ordem de citação, fixando o termo “Biblioteconomia” como significante principal na representação da informação. O termo “história”, neste exemplo, se apresenta como um delimitador que, a um só tempo, promove a especificidade na representação e a probabilidade de precisão na recuperação da informação. Em outras palavras, significa dizer que o documento analisado e representado não trata do conteúdo “História”, mas da “Biblioteconomia”, a partir de uma abordagem histórica. Essas especificações realizadas nos dois exemplos anteriores (Quadros 1 e 2), por intermédio do acréscimo de delimitadores, são



imprescindíveis nos sistemas de recuperação de informação porque estão diretamente relacionadas à identificação da relevância da informação pelos usuários que a demandam e a recuperam.

Nos sistemas que adotam a linguagem pré-coordenada, o primeiro termo é o determinante da recuperação da informação, sendo, por conseguinte, necessário o uso de entradas múltiplas ou de diversas remissivas para explicitar todos os conceitos relevantes.

Exemplo 1:

Enunciado: Implantação de bibliotecas no Brasil no período colonial

Entradas múltiplas:

BIBLIOTECAS – IMPLANTAÇÃO – BRASIL COLONIAL

BRASIL COLONIAL – BIBLIOTECAS – IMPLANTAÇÃO

Exemplo 2:

Enunciado: Sociologia e Ciência Política

Entradas múltiplas:

CIÊNCIA POLÍTICA – SOCIOLOGIA

SOCIOLOGIA – CIÊNCIA POLÍTICA

Diante disso, deve-se observar que a adoção de linguagens pré-coordenadas, com o uso de diferentes entradas e diversas remissivas, resulta em alto custo direto e indireto para o sistema de recuperação da informação. Nesse conjunto, pode-se destacar o custo com pessoal, fichas ou registros de assunto, espaço nos catálogos ou índices impressos e nas bases de dados, impressão, manutenção dos arquivos e tempo na busca de informação.

As **linguagens pós-coordenadas** são aquelas que combinam ou coordenam os termos ou símbolos que representam os conteúdos dos documentos no momento da recuperação da informação. Isso significa dizer que a representação da informação é realizada usando termos simples, na medida em que a coordenação entre estes é realizada em outro momento. Assim, o processo de coordenação se dá na elaboração das estratégias de busca e é realizada pelo usuário ou, exclusivamente no caso de buscas delegadas, pelo bibliotecário.

Essas linguagens são usadas preferencialmente em sistemas de recuperação de informação automatizados e, por isso, são chamadas de linguagens modernas (CESARINO; PINTO, 1978). Como exemplo, pode-se citar o tesouro.

Os sistemas pós-coordenados trabalham preferencialmente com conceitos simples, lembrando-se que estes podem ser formado por uma ou mais palavras.

Exemplos: Biblioteca (**uma palavra**).

Biblioteca universitária (**duas palavras**).

Busca delegada – A busca delegada, como a própria expressão designa, corresponde a uma transferência de responsabilidade entre usuário e bibliotecário. Na busca delegada, o usuário, em função de uma série de fatores, dentre os quais se pode destacar a dificuldade na elaboração das estratégias de busca, solicita ao bibliotecário a realização do conjunto de procedimentos que visa à procura, à identificação e à localização, no catálogo impresso ou on-line, do documento que responde a uma necessidade de informação específica.

Fonte: CARNEIRO, M. V. Diretrizes para uma política de indexação. **Revista da Escola de Biblioteconomia da UFMG**, Belo Horizonte, v. 14, n. 2, p. 221-241, set. 1985. Disponível em: <<http://portaldeperiodicos.eci.ufmg.br/reb/>>. Acesso em: 26 dez. 2015.

É necessário compreender que, embora haja duas palavras na entrada do segundo exemplo, trata-se de um conceito simples. Você precisa observar que “biblioteca universitária” não designa dois conceitos simples (**biblioteca** e **universidade**), que formam um conceito composto.

Diferentemente, os exemplos abaixo são formados por duas palavras, que designam dois conceitos simples e formam um conceito composto.

Exemplos: Planejamento de biblioteca.

Recuperação da informação.

Trata-se de conceitos simples (**planejamento** e **biblioteca**, bem como **recuperação** e **informação**) que, coordenados dão origem a conceitos compostos.

A partir da complexidade de análise e síntese, sobretudo, considerando as dispersões semânticas e sintáticas, observa-se a necessidade de definição de um conjunto de estratégias que auxiliem o analista nos processos de identificação, interpretação e seleção dos conceitos objetos da análise da informação.



2.2 ESTRATÉGIAS DE LEITURA

Preliminarmente, você deverá considerar que a realização dos processos de análise e síntese da informação tem como referências básicas o documento, que é a base material, o conteúdo informacional e os procedimentos adotados. Nesse conjunto, merece destaque a leitura, que possibilita a interpretação, a extração de conceitos e a síntese do conteúdo informacional.

Segundo *Dias e Naves (2007)*, existem diferentes tipos de textos, conforme o enfoque, a finalidade e os procedimentos adotados. Nesse universo, esses autores destacam as classificações de *Egon Werlich* e *Ernst Ulrich Grosse*. Na sua classificação, *Werlich (1975)* considera o contexto extralinguístico, sobretudo social, e a estrutura das orações: argumentativa, descritiva, expositiva (analítica e sintética), instrutiva e narrativa. *Grosse (1976)*, por sua vez, considera as seguintes funções: automanifestação (diários, autobiografia etc), indicativa de contato (saudações, felicitações, etc.), indicativa de grupo (músicas religiosas, canções, etc.), normativa (estatutos, leis, normas, etc.), ordem/exigência (solicitações, questionamentos, etc.), poética (contos, novelas, poemas etc) e transferência de informação (correspondências, informes, notícias, etc.). Os tipos de textos mais comuns, contudo, são os narrativos e os informativos.

Na definição de estratégias de leituras a serem adotadas na análise da informação, deve-se considerar a eficiência destas na própria análise e também nas estratégias de busca que poderão vir a ser realizadas no processo de recuperação da informação. Essa assertiva tem por base o entendimento de que os textos são produzidos a partir da articulação de conceitos e conteúdos formando o que se denomina esquema.

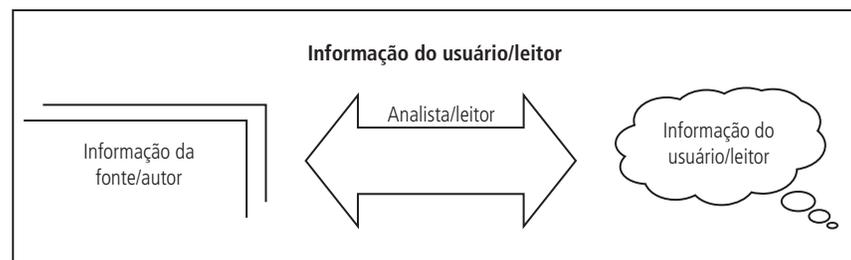
Conforme Kato (1985), há dois tipos de estratégias de leitura: **cognitivas** e **metacognitivas**. As primeiras são realizadas de forma inconsciente e automática pelo leitor. As segundas se dão a partir de comportamentos conscientes e, portanto, controlados do leitor. A autora considera, contudo, que o texto inteligível corresponde àquele que requer a aplicação equilibrada destes dois tipos de estratégias. No domínio destas, há possibilidade de definição de estratégias particulares, considerando a qualidade do texto e o conhecimento prévio do analista.

Outra questão que merece destaque inicialmente diz respeito à impossibilidade de o analista realizar a leitura palavra por palavra dos documentos que se encontram em processo de tratamento. A partir desta constatação, é forçoso considerar que, na análise da informação, torna-se necessário o estabelecimento de um conjunto de eficientes estratégias de leitura que possibilite a interpretação e a seleção de conceitos, sem a realização da leitura integral do texto-fonte.

Não se pode perder de vista também que o bibliotecário que executa o processo de análise da informação, na maioria das vezes, não é um especialista no conteúdo dos documentos que estão sendo analisados e representados. Há indubitavelmente a necessidade manifesta do estabelecimento de estratégias de leitura que facilitem esse processo.

A Figura 10 evidencia que as estratégias de leitura devem ser elaboradas considerando dois universos de referência, a saber, o contexto de produção (informação da fonte/autor) e o contexto de uso (informação do usuário/leitor) dos conteúdos informativos.

Figura 10 – Esquema geral de estratégia de leitura



Fonte: Produção do próprio autor (2015).

O conteúdo informacional compreende o conjunto de estruturas significantes, que, interpretadas e assimiladas, transformam-se em informação. O conteúdo é, em outros termos, composto por unidades informacionais. Nesse contexto, essas unidades se referem à materialidade imediata da análise da informação, que se pode encontrar nos

formatos de texto, som, imagem ou em combinações diversas entre esses formatos.

Por outro lado, é importante observar que, conforme *Cintra* (1987), na leitura documentária, há uma ruptura do princípio de cooperação entre autor-leitor/analista, na medida em que não existe previsão, por parte do autor, da ação do analista como leitor. A fonte/autor é pensada a partir da relação imediata com o usuário/leitor.

A seleção de estratégias de leitura, de acordo com *Cintra* (1987), depende de uma série de condições, dentre as quais se destacam:

- a) a finalidade da leitura;
- b) a experiência do leitor;
- c) a maturidade na tarefa de ler;
- d) o tipo de texto objeto da leitura;
- e) a concentração da atenção em partes específicas do texto;
- f) o grau de novidade para o leitor e
- g) a motivação para a realização da leitura.

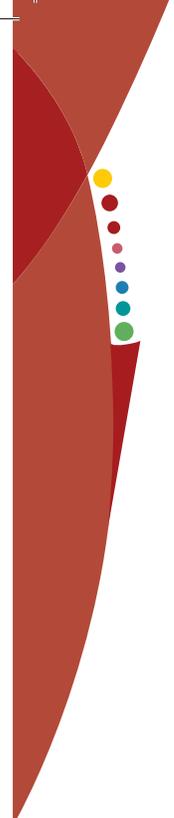
Observe que as estratégias de leituras estão relacionadas ao texto, ao leitor e ao usuário do sistema de informação. As condições compostas por experiência, maturidade, novidade e motivação se referem ao estado de preparo e de espírito em que se encontra o analista da informação em relação ao conteúdo objeto de análise e representação. As demais variáveis, isto é, a finalidade da leitura, o tipo de texto e a concentração em partes textuais, estão relacionadas ao texto em si e ao interesse do usuário pelos conteúdos específicos do documento. O analista é forçado a compreender e considerar a dinâmica da organização do texto, na relação com a finalidade de sua inclusão no sistema de recuperação da informação.

A primeira estratégia de leitura é o reconhecimento de que o texto é formado por um conjunto de estruturas significantes, mas que se encontram fora do seu contexto de produção. De modo geral, o texto como base de leitura apresenta as seguintes características:

- a) unidade complexa e dinâmica de significação;
- b) sinal aberto de significação;
- c) intermediação entre autor (conteúdo) e usuário (necessidade/demanda de informação).

Conforme *Dias e Naves* (2007, p. 37): "A dificuldade de uma descrição científica do texto deriva do fato de ele ser um sinal *open ended*, tendo um significado dinâmico e aberto, sobre o que o documentalista deveria pensar em grande profundidade". Assim, dada a sua complexidade constitutiva, é fundamental considerar que o texto apresenta uma estrutura geral e algumas particularidades que se relacionam à sua tipologia. A estrutura está afeita ao arranjo organizacional e à apresentação das informações.





Além do conhecimento prévio dos tipos de texto, outra estratégia é compreender as estruturas textuais, que possibilitam diferentes tipos de leituras:

- a) **conhecimento da microestrutura:** exige a realização da leitura palavra por palavra;
- b) **conhecimento da macroestrutura:**
 - permite a compreensão das relações entre as partes que compõem a unidade textual e
 - promove a leitura de tópicos e a compreensão das respectivas estruturas sintáticas e semânticas;
- c) **conhecimento da superestrutura:**
 - reduz a necessidade de apoio na leitura palavra por palavra;
 - diminui a dificuldade de integração das informações que compõem o texto como um todo e
 - facilita a identificação das ideias centrais do texto, uma vez que tem como referência a identificação dos elementos básicos constituintes (CINTRA, 1987).

É importante ter uma postura crítica e reconhecer que as circunstâncias, o tipo de texto, o conhecimento das estruturas textuais e conhecimentos prévios sobre o conteúdo informativo são elementos fundamentais na construção das estratégias de leituras. A análise pode resultar na integração de diferentes estratégias.

Um recurso bastante utilizado na análise e na síntese da informação se refere à denominada leitura técnica, que, nas palavras de Lancaster (2004, p. 24), corresponde a “[...] um misto de ler e ‘passar os olhos’ no texto”. Esse tipo de leitura centra basicamente nas seguintes partes do documento:

- a) título;
- b) subtítulo, se houver;
- c) sumário;
- d) resumo, se houver;
- e) prefácio, se houver;
- f) apresentação, se houver;
- g) introdução;
- h) títulos dos capítulos e
- i) referências (bibliográficas).

É relevante tomar os devidos cuidados na realização de leituras técnicas, sobretudo, no título, no subtítulo e nos títulos de capítulos do documento, e atentar para o uso de figuras de linguagem, especialmente metáforas e metonímias. Essas construções textuais são bastantes presentes em textos de ficção ou literários. Assim, se necessário, torna-se imprescindível acrescentar àqueles elementos a leitura dos primeiros parágrafos dos principais itens (capítulos e subcapítulos) do documento.

2.3 INSTRUMENTOS AUXILIARES E PRODUTOS

Para cada processo de análise e síntese, temática e descritiva, da informação existem, pelo menos, um correspondente instrumento auxiliar e um produto como resultado. Os instrumentos auxiliares podem ser usados no processo de análise da informação, mas são, muitas vezes, direcionados especificamente ao processo de tradução, isto é, à passagem da linguagem natural para a linguagem do sistema de informação. Assim, torna-se importante neste momento ter um mínimo de compreensão acerca dos principais instrumentos auxiliares e produtos.

As linguagens de informação, segundo o arranjo, podem ser **classificadas** ou **alfabéticas** e têm finalidades bastante específicas no sistema de recuperação da informação. Como tais, apresentam características bastante diversas, considerando, sobretudo, os arranjos, os métodos utilizados e os tipos de controle de linguagem a que se destinam. Com efeito, todos esses instrumentos auxiliares têm por finalidade o controle, sintático e/ou semântico, no processo representacional.

As **classificações bibliográficas**, como a própria expressão designa, são linguagens classificadas utilizadas para representar tematicamente os conteúdos informativos dos documentos, para organizá-los e arranjá-los nas estantes e para possibilitar a sua posterior localização. *Langridge* (2006) esclarece que o sistema de classificação bibliográfica é o mapeamento dos conceitos e de suas relações semânticas em qualquer área do conhecimento. O processo de classificação, portanto, não se refere à simples atribuição de notações, mas essencialmente à análise do conteúdo do documento e à decisão sobre a classe ou subclasse em que estas se encontram no universo do conhecimento. Esse tipo de linguagem é caracterizado, principalmente, pelo seu arranjo sistemático, considerando as relações de coordenação e subordinação/superordenação existentes entre os conteúdos que o compõem. Nesse universo, destacam-se a CDD e a CDU, que classificam a unidade do conhecimento em 10 (dez) grandes classes.

As classificações bibliográficas utilizam algarismo e/ou símbolos para representar os conteúdos dos documentos. Essa representação é chamada de notação, que, conforme a sua formação, pode ser descrita como pura ou mista. Para *Piedade* (1983), a notação pura é composta por caracteres de um único tipo, números ou letras, por exemplo. A notação mista, por sua vez, é composta por mais de um tipo de caracter. São exemplos desses tipos de notação, respectivamente, a CDD e a CDU.



Quadro 3 – Classificações bibliográficas

CDD	CDU
História da Biblioteconomia => 020.9	História da Biblioteconomia => 02(091)

Fonte: Produção do próprio autor (2015).

Observe, no Quadro 3, que, na notação da CDD, constam apenas números, na medida em que o “ponto” não tem qualquer significado, servindo apenas para facilitar a leitura. Por outro lado, na notação da CDU, constam números e símbolos, no caso em tela, os parêntesis, que designam os números auxiliares comuns de forma de apresentação do conteúdo do documento.

Para melhor compreensão da importância da análise da informação no processo classificatório, você pode observar, nas Figuras 11 e 12, que a estrutura da CDD é composta por assuntos principais e assuntos auxiliares, que compõem, respectivamente, os esquemas (000-900) e as tabelas auxiliares (T1-T6). A representação de um conteúdo informativo pode ser realizada a partir da coordenação entre assunto principal e assuntos auxiliares, conforme necessidades e disposições estabelecidas no sistema de classificação.

Figura 11 – Esquemas da CDD 22ª edição

CDD	
000	Obras Gerais
100	Filosofia, Fenômenos Paranormais e Psicologia
200	Religião
300	Ciências Sociais
400	Língua
500	Ciências Naturais e Matemática
600	Tecnologia (Ciências Aplicadas)
700	Belas Artes
800	Literatura
900	Geografia, História e Ciências Auxiliares

Fonte: CDD 22ª edição.

Assim, como bem alerta *Lancaster* (2004), a classificação e o uso das classificações bibliográficas não podem ser entendidos como mera atribuição de números ou, em outros termos, a realização da segunda etapa da representação da informação, a tradução. Ao contrário, trata-se, essencialmente, de analisar os conteúdos informacionais e decidir a classificação e o conjunto de relações que existe entre estes e os demais que compõem o acervo do sistema de informação. Então, observa-se que, além de classificar o conteúdo informacional, é indispensável considerar que será exigida do analista a competência interpretativa e sintética para compreender sobre o que trata o documento e quais são os seus assuntos principais e secundários.

Além das 10 classes gerais, denominadas de esquemas, a CDD é composta por 6 (seis) tabelas auxiliares, conforme Figura 12.

Figura 12 – Tabelas auxiliares da CDD 22ª edição

Tabela 1	Subdivisão Padrão
Tabela 2	Áreas Geográficas, Períodos Históricos e Pessoas
Tabela 3	Subdivisões para Artes, para Literatura Individual, para formas Literárias Específicas
Tabela 3A	Subdivisões para trabalhos de ou sobre Autores Individuais
Tabela 3B	Subdivisões para trabalhos de ou sobre mais de um Autor
Tabela 3C	Notação para ser acrescentada com instrução na Tabela 3B, 700.4, 791.4, 808-809
Tabela 4	Subdivisões de Línguas Individuais e Famílias de Línguas
Tabela 5	Etnia, Grupos Nacionais
Tabela 6	Linguagem

Fonte: CDD 22ª edição.

Para melhor compreensão, pode-se considerar hipoteticamente que o bibliotecário decidiu que um determinado documento deve ser classificado em **020.905**. Essa não é uma decisão arbitrária, mas resulta do processo classificatório, que tem na sua base a análise do conteúdo informativo do referido documento e, complementarmente, a análise das temáticas disponíveis no sistema de classificação, no caso em tela na CDD.

Primeiro, o bibliotecário identificou que o documento tem como conteúdo “História da Biblioteconomia no século XXI”, conforme Quadro 4. Segundo, que existe possibilidade de classificar usando os esquemas e as tabelas auxiliares da CDD, conforme Quadro 5. Essa compreensão se deve às análises realizadas, respectivamente, no texto-fonte e no sistema de classificação. Com efeito, no contexto desta disciplina, basta saber que, no exemplo, tem-se um assunto composto formado por três assuntos simples, a saber, “História”, “Biblioteconomia” e “século XXI”.

Exemplo:

Quadro 4 – Análise e síntese do conteúdo do documento

Categoria analítica	Temática
Conteúdo principal	Biblioteconomia
Conteúdo secundário 1	História
Conteúdo secundário 2	Século XXI
Conteúdo-síntese	História da Biblioteconomia no século XXI

Fonte: Produção do próprio autor (2015).



Quadro 5 – Análise e síntese do conteúdo usando a CDD

Categoria analítica	Temática	Notação
Conteúdo principal	Biblioteconomia	020
Conteúdo secundário 1	História	-09
Conteúdo secundário 2	Século XXI	-0905
Conteúdo-síntese	História da Biblioteconomia no século XXI	020.905

Fonte: Produção do próprio autor (2015).

O número de classificação ou notação acima resulta, portanto, das atividades de análise do conteúdo e tradução. Mais que isso, cada processo classificatório exige análises com extensões e combinações, por vezes, bastante diversas, que estão condicionadas ao conteúdo do documento, à finalidade da incorporação deste ao acervo e à relação com as necessidades e demandas dos usuários.

A **lista de cabeçalhos de assunto** é outro importante instrumento usado no controle de vocabulário em sistemas de recuperação da informação. Como tal, esse instrumento é imprescindível ao processo de tradução, que, conforme visto anteriormente, acontece em dois momentos no sistema de recuperação da informação, na representação da informação (tratamento da informação) e na elaboração de estratégias de busca (recuperação da informação).

Diferentemente das classificações bibliográficas, a lista de cabeçalhos de assunto faz uso da própria linguagem natural e, por conseguinte, tem arranjo alfabético. Esse instrumento de controle de vocabulário se refere a um conjunto de palavras usado para expressar da forma mais adequada possível o assunto dos documentos. Do ponto de vista estrutural, este instrumento é uma lista de assuntos organizada alfabeticamente, estabelecendo relações de equivalência, associativas e hierárquicas.

Conforme *Lancaster* (2004), além de não distinguir claramente as relações associativas e hierárquicas, nas listas de cabeçalhos de assunto, define-se relações hierárquicas imperfeitas, como se observa a seguir na Figura 13, no conjunto de remissivas dos tipos **ver** (x) e **ver também** (xx). Observe que, no arranjo do cabeçalho de assunto, essas remissivas se apresentam como uma verdadeira rede de referências cruzadas que se expande em toda a lista de cabeçalhos.

Figura 13 – Lista de cabeçalhos de assunto

ARQUIVO xx Biblioteca Centro de informação Museu Teoria das três idades	BIBLIOTECA PÚBLICA xx Biblioteca escolar Biblioteca especial Biblioteca especializada Biblioteca universitária
ARQUIVOLOGIA xx Biblioteconomia Ciência da Informação Documentação Museologia	BIBLIOTECA UNIVERSITÁRIA xx Biblioteca escolar Biblioteca especial Biblioteca especializada Biblioteca pública
BIBLIOTECA x Biblioteca escolar Biblioteca especial Biblioteca especializada Biblioteca pública Biblioteca universitária	BIBLIOTECONOMIA xx Arquivologia Ciência da Informação Documentação Museologia
BIBLIOTECA ESCOLAR xx Biblioteca especial Biblioteca especializada Biblioteca pública Biblioteca universitária	CIÊNCIA DA INFORMAÇÃO xx Arquivologia Biblioteconomia Documentação Museologia
BIBLIOTECA ESPECIAL xx Biblioteca escolar Biblioteca especializada Biblioteca pública Biblioteca universitária	CIENTISTA DA INFORMAÇÃO xx Bibliotecário
BIBLIOTECA ESPECIALIZADA xx Biblioteca escolar Biblioteca especial Biblioteca pública Biblioteca universitária	[...]
	ZONA RURAL x Perímetro rural
	xx Perímetro urbano
	ZONA URBANA x Perímetro urbano
	xx Perímetro rural

Fonte: Produção do próprio autor (2015).

Como visto anteriormente, as remissivas do tipo **ver** têm função prescritiva e, por isso, estabelece o controle de vocabulário definindo os termos preferenciais e restringindo o uso da linguagem. Trata-se de definir termos substitutivos para aqueles não autorizados. As remissivas do tipo **ver também** têm função sugestiva, realizam o controle de vocabulário definindo termos alternativos e, neste caso, ampliam as possibilidades do uso da linguagem. Trata-se de definir termos alternativos para aqueles autorizados.

Assim como acontece com as classificações bibliográficas, os cabeçalhos de assunto também são utilizados para fazer o controle sintático, nas situações em que estabelecem o uso da ordem de citação de entradas



compostas. Trata-se, em outros termos, de definir, em assuntos compostos e complexos, o cabeçalho e seu(s) respectivo(s) subcabeçalho(s).

Nas análises e definições de subcabeçalhos de assunto, é imprescindível considerar que, em decorrência do princípio da pré-coordenação, estes servem tão-somente para especificar a representação e precisar melhor a pesquisa, não se constituindo em pontos de acesso em função da unidimensionalidade que o caracteriza. As buscas são realizadas sempre pelos cabeçalhos, na medida em que, em um primeiro momento, os subcabeçalhos se encontram ocultos na estrutura sintática. Por isso, na definição da prioridade dos termos, por vezes, torna-se necessário utilizar a denominada entrada invertida – faz-se uma inversão da ordem dos termos que se encontram no enunciado original do conteúdo.

Exemplos:

História da Biblioteconomia ver BIBLIOTECONOMIA – HISTÓRIA	Teorias da classificação ver CLASSIFICAÇÃO – TEORIA	Tipos de resumos ver RESUMOS – TIPOS
--	---	---

A ordem de citação exige a definição de prioridades entre dois ou mais assuntos. Essa decisão se reveste, em grande medida, da subjetividade do analista. Buscando, pelo menos, minimizar esta dificuldade na definição das relações sintáticas entre os assuntos de um enunciado, alguns estudiosos procuraram definir fundamentos que dessem maior consistência a esta atividade.

Julius Otto Kaiser, por exemplo, sugeriu, em 1911, que, na ordem de citação de assuntos compostos, deve-se seguir a seguinte lógica: concreto – processo.

Exemplo:

Enunciado: Administração de biblioteca

Ordem de citação: BIBLIOTECA – ADMINISTRAÇÃO.

Já em 1960, E. J. Coates, refletindo sobre a lógica definida por Kaiser, propôs a seguinte ordem de citação: coisa – parte – material – ação.

Exemplo:

Enunciado: O uso de bases de dados no serviço de referência da biblioteca pública

Ordem de citação: BIBLIOTECA PÚBLICA – SERVIÇO DE REFERÊNCIA – BASES DE DADOS – USO.

A ordem de citação pode ser realizada usando parte das categorias, conforme exigência do enunciado.

Exemplo:

Enunciado: Estantes de bibliotecas

Ordem de citação: BIBLIOTECAS – ESTANTES.

A definição de ordem de citação mais conhecida, contudo, se refere às cinco categorias propostas por S. R. Ranganathan, em 1933, personalidade, matéria, energia, espaço e tempo, mais conhecidas como PMEST, isto é: P (**Personality**), M (**Matter**), E (**Energy**), S (**Space**) e T (**Time**).

Exemplo:

Enunciado: Implantação de bibliotecas escolares no Brasil no século XIX

Ordem de citação: BIBLIOTECAS ESCOLARES – IMPLANTAÇÃO – BRASIL – SÉCULO XIX.

Embora tenha havido melhoras na elaboração das listas de cabeçalhos de assunto, o seu uso exige do analista da informação uma postura bastante crítica, sobretudo porque apresentam algumas fragilidades, tais como:

- a) o nível de especificidade não é suficiente para muitos documentos, principalmente para os assuntos cada vez mais complexos;
- b) os cabeçalhos de assuntos não são elaborados sistematicamente;
- c) as referências não são elaboradas sistematicamente e, quando construídas, centram fundamentalmente nas remissivas do geral para o específico;
- d) os subcabeçalhos não são arranjados de forma homogênea (CESARINO; PINTO, 1978).

Assim, trata-se de considerar a seguinte observação realizada há alguns anos por *Cesarino e Pinto* (1978, p. 284), mas bastante atual:

O que se pode dizer a respeito das listas de cabeçalhos de assunto é que, ao serem adotadas por uma determinada biblioteca, devem ser adaptadas a ela. Não se deve exigir nenhuma rigidez quanto ao emprego de cabeçalhos, e a adaptação desses e a adoção de novos deve ser uma prática normal num serviço de biblioteca.

Nos sistemas pós-coordenados, adotam-se conceitos simples, uma vez que a formação de conceitos compostos e complexos é realizada por intermédio da coordenação, no momento da busca. O **tesauro** é o principal instrumento utilizado no controle de vocabulário em sistemas de recuperação da informação que adotam a pós-coordenação. Esse tipo de controle tem como principal arranjo o alfabético, mas intrinsecamente também faz uso do arranjo hierárquico, ao estabelecer as relações semânticas entre os descritores.

Em se tratando do tesauro, que é usado principalmente em sistemas automatizados, a definição da ordem de citação dos termos que formam assuntos compostos e complexos não representa maiores problemas para o analista, uma vez que a coordenação dos termos é realizada pelo usuário no momento da busca. Esse tipo de vocabulário trabalha com conceitos individuais e, por conseguinte, é formado por cabeçalhos simples denominados de “descritores” ou “termos”.

Assim, no que se refere à construção de linguagens documentárias, a unidade informacional é o conceito, embora, é claro, ela deva estar expressa por termos, simbolicamente representados por palavras. Desta forma, os conceitos podem estar representados por apenas uma palavra, ou por uma composição de palavras, desde que esta composição designe um único conceito (DODEBEI, 2002, p. 79).



Exemplos:

COMPUTAÇÃO	COMUNICAÇÃO SOCIAL	PESQUISAS EM BIBLIOTECONOMIA
------------	--------------------	------------------------------

Os exemplos acima têm composições diversas, sendo que o primeiro e o segundo descrevem, cada um deles, em que pese a quantidade distinta de palavras, respectivamente, uma e duas, representa um único conceito – “Computação” e “Comunicação Social”. O terceiro também é composto por duas palavras, mas, ao contrário do segundo exemplo, representa dois conceitos, “Pesquisa” e “Biblioteconomia”. Apenas neste caso, tem-se a situação de dois descritores e, portanto, de duas entradas, tanto no tesouro quanto no sistema de recuperação da informação. O conceito deve ser entendido, então, como unidade de conhecimento.

Figura 14 – Estrutura de relações semânticas no tesouro

ARQUIVO TE Teoria das três idades TR Arquivologia Biblioteca Centro de informação Museu	BIBLIOTECA PÚBLICA TG Biblioteca TR Biblioteca Escolar Biblioteca Especial Biblioteca Especializada Biblioteca Universitária
ARQUIVOLOGIA TR Arquivo Biblioteconomia Ciência da Informação Documentação Museologia	BIBLIOTECA UNIVERSITÁRIA TG Biblioteca TR Biblioteca Escolar Biblioteca Especial Biblioteca Especializada Biblioteca Pública
BIBLIOTECA TE Biblioteca Escolar Biblioteca Especial Biblioteca Especializada Biblioteca Pública Biblioteca Universitária TR Arquivo Centro de Informação Museu	BIBLIOTECONOMIA TR Arquivologia Ciência da Informação Documentação Museologia
BIBLIOTECA ESCOLAR TG Biblioteca TR Biblioteca Especial Biblioteca Especializada Biblioteca Pública Biblioteca Universitária	CENTRO DE INFORMAÇÃO TR Arquivo Biblioteca Museu
	CIÊNCIA DA INFORMAÇÃO TR Arquivologia Biblioteconomia Documentação Museologia

BIBLIOTECA ESPECIAL	CIENTISTA DA INFORMAÇÃO
TG	TE
Biblioteca	Bibliotecário
TR	
Biblioteca Escolar	[...]
Biblioteca Especializada	
Biblioteca Pública	ZONA RURAL
Biblioteca Universitária	UP
	Perímetro rural
BIBLIOTECA ESPECIALIZADA	TR
TG	Perímetro urbano
Biblioteca	Zona urbana
TR	
Biblioteca Escolar	ZONA URBANA
Biblioteca Especial	UP
Biblioteca Pública	Perímetro Urbano
Biblioteca Universitária	TR
	Perímetro rural
	Zona Rural

Fonte: Produção do próprio autor (2015).



Ao trabalhar com conceitos, além de controlar as dispersões semânticas e sintáticas, esse instrumento define, conforme esclarecem *Cintra et al.* (2002), um domínio conceitual formado pelos conceitos e pelas relações semânticas e sintáticas entre estes. Mais que isso, esse domínio conceitual ou, de forma mais precisa, terminológico aproxima dois universos o dos documentos e o das necessidades dos usuários. É nesses domínios que a elaboração desse tipo de instrumento tem como fundamento, respectivamente, a garantia literária e a garantia do usuário.

A garantia literária tem como princípio fundamental o entendimento de que, na definição do domínio e das classes que compõem uma linguagem de indexação, não se deve tomar como referências as classificações do conhecimento, mas a literatura da área objeto de análise e síntese (DODEBEI, 2002). Assim, um vocabulário controlado (cabeçalho de assunto ou tesouro) usado, por exemplo, no processo de representação dos conteúdos de documentos da Ciência da Informação deve ser elaborado a partir da produção científica desta área de conhecimento. Trata-se, nas palavras de *Cesarino e Pinto* (1978), de elaborá-lo a partir de uma abordagem analítica.

A garantia do usuário, por sua vez, tem por base o princípio de uso, segundo o qual a linguagem adotada no sistema de recuperação de informação deve considerar a conveniência, as necessidades e as demandas dos usuários deste sistema. Uma linguagem que tem como referência este tipo de garantia também deve ser elaborada adotando uma abordagem analítica (CESARINO; PINTO, 1978). Na prática, uma linguagem adotada na representação do conteúdo dos documentos, em uma biblioteca escolar, deve tomar como referência parâmetros da comunidade escolar específica.

O tesouro é o instrumento de controle de vocabulário que permite a combinação das relações da linguagem documentária com os sistemas de buscas automatizados. Uma das suas particularidades é justamente o fato de trabalhar com conceitos simples e permitir o uso de lógicas matemáticas na elaboração de estratégias de busca, condicionando melhora a comunicação entre o sistema de recuperação da informação e os usuários.

Do ponto de vista do arranjo, como se observa no extrato apresentado na Figura 14, o tesouro apresenta duas categorias de relação: coordenação e subordinação/superordenação. Por outro lado, percebe-se também que as relações de subordinação são mais claras em função de o arranjo deste instrumento expressar as relações hierárquicas de forma explícita, por meio das remissivas “Termo Específico (TE)” e “Termo Genérico (TG)”. As relações hierárquicas compreendem relações “gênero x espécie” e “todo x parte”.

Além das relações de coordenação (equivalência, associativas e relacionadas) e de subordinação/superordenação (hierárquicas), o tesouro adota “Notas Explicativas (NE)”, que servem para esclarecer o alcance conceitual do descritor.

Exemplos:

CIÊNCIA DA INFORMAÇÃO	COMUNICAÇÃO
NE Campo de conhecimento dedicado aos estudos dos atributos da informação e dos respectivos processos, metodologias e tecnologias.	NE Campo de conhecimento. Inclui Jornalismo, Publicidade e Propaganda, Radialismo e Relações Públicas.

Como resultado dos processos de análise e síntese (catalogação, classificação, indexação e elaboração de resumo), existe uma infinidade de produtos que vêm sendo elaborados no decorrer da história dos registros do conhecimento, desde as primeiras listas de inventários aos mais atuais e sofisticados repertórios de informação. Assim, podem-se citar catálogos, bibliografias, bibliografia de bibliografias, índices, inventários, metadados e resumos, entre outros. Lembrando-se que, com o uso das tecnologias digitais, começam a ser planejados e implementados produtos que, na sua essência, resultam de combinações de alguns já existentes.

Os produtos obtidos pelos processos de representação documentária, em seus vários graus de redução semântica, quer sejam: o texto ou objeto na íntegra, resumos, referências bibliográficas, índices temáticos ou notações de classe sob as formas de codificações textuais, imagéticas, sonoras e digitais, organizadas em inventários, catálogos, repertórios, índices, bibliografias, vão ser disseminados à sociedade em função da sua demanda por informação (DODEBEI, 2002, p. 36).

Nesse conjunto, contudo, torna-se mais fácil a compreensão daqueles produtos que tem relação mais direta com os instrumentos auxiliares acima estudados. Essa abordagem facilita inclusive a compreensão de que não se pode confundir linguagem documentária ou linguagem de informação com os produtos resultantes da representação da informação realizada com o seu respectivo auxílio.

No que se refere aos produtos da análise e síntese da informação, três se destacam: o catálogo, o índice e o resumo.

O **catálogo** é o mais tradicional produto dos processos de análise e síntese da informação em função da sua longa história nos registros do conhecimento. Tradicionalmente, é conceituado como um conjunto de fichas catalográficas (7,5 cm x 12,5 cm) organizadas de acordo com um

princípio e a finalidade. Ele tem um papel fundamental na organização e gestão das coleções, pois é um importante instrumento de inventário e controle. Isso implica dizer que este produto tem como objeto de análise e síntese os documentos que se encontram disponíveis na biblioteca, possibilitando a visibilidade do seu conteúdo informacional.

De modo mais preciso, o catálogo é o principal produto da organização da informação em uma biblioteca, uma vez que é o responsável pela organização das informações documentais, descritivas e temáticas, e pela recuperação das informações que compõem os conteúdos dos diversos documentos disponíveis no acervo da biblioteca.

Observa-se que, no seu conjunto, este produto é resultado das atividades de classificação, catalogação e indexação. Na Figura 2 (ver p. 20), por exemplo, você pode identificar, além da catalogação (elementos descritivos no centro da ficha catalográfica), a classificação (02 "CDU" e 020 "CDD") e a indexação (1. Ciência da Informação).

Considerando os pontos de acesso que servem de parâmetro para organização das fichas, há três tipos de catálogos, a saber: catálogo de autor (ficha principal – Figura 2), catálogo de assunto e catálogo de título.

De modo geral, esses catálogos são ordenados alfabeticamente, mas podem também adotar na sua organização algum tipo de sistema de classificação como, por exemplo, as classificações bibliográficas. Assim, conforme a forma de arranjo, os catálogos são classificados em alfabético ou sistemático.

Por fim, os catálogos podem ser impressos ou digitais. Os catálogos digitais apresentam, em grande medida, as mesmas informações documentais presentes nos catálogos impressos e seguem os padrões estabelecidos para estes. Nos documentos disponíveis na **internet**, contudo, as informações descritivas ganham outro formato, sendo denominadas genericamente de metadados, dados sobre dados. *Modesto* (2005) define metadados como conjunto de elementos que possui uma semântica padronizada, possibilitando descrever recursos eletrônicos de maneira bibliográfica.

O índice é outro produto da análise e síntese da informação e resulta do processo de indexação. Trata-se, portanto, de um tipo de representação temática, que também tem uma longa tradição, especialmente em instituições especializadas que realizam os **serviços de indexação e resumo**.

Segundo a NBR 6034/ABNT (2004), que estabelece os requisitos de apresentação e os critérios básicos de elaboração de índices, eles podem ser classificados de acordo com dois critérios, a saber: a ordenação e o enfoque. Segundo a ordenação, o índice pode ser **alfabético, alfanumérico, cronológico, numérico** e **sistemático**. Quanto ao enfoque, o índice pode ser **especial** ou **geral**. O índice especial compreende ainda subcategorias, conforme a sua materialidade: **assuntos, autores, anunciantes e matérias publicitárias, citações, nomes geográficos, pessoas e/ou entidades** e **títulos**. O índice geral corresponde àquele que combina duas ou mais categorias, tais como assuntos e títulos.

Na prática, além da combinação no domínio do enfoque, podem-se ter outras categorias híbridas, considerando a ordenação e o enfoque. Assim, podem-se ter índice alfabético de assunto, índice sistemático de assunto, índice alfabético de autores, índice alfabético de títulos, entre outros.



Esse tipo de produto pode fazer parte do próprio documento a que representa ou ser produzido de forma avulsa. Quando produzido na própria fonte como, por exemplo, em livros, o índice corresponde a uma lista de termos organizados em ordem alfabética, seguidos do(s) número(s) da(s) página(s) em que se encontra a informação indexada. Embora a característica principal de todos os índices seja auxiliar no processo de localização da informação, este produto é denominado de **índice remissivo**. Os índices mais sofisticados apresentam na sua organização relações associativas (*ver* e *ver também*) e hierárquicas entre os termos, sendo que essas são quase sempre evidenciadas por simples recuos. Segue, na Figura 15, representação de extrato de índice alfabético de assunto produzido na própria fonte.

Figura 15 – Índice alfabético de assunto

A
Arquivo, 5, 8, 15, 20, 25, 36
Ver também Biblioteca
Arquivo corrente, 15, 20
Arquivo intermediário, 20, 25
Arquivo permanente, 28, 36
Arquivo privado, 8
Arquivo público, 5
Arquivologia, 1-3, 8, 22, 35
Cursos, 21
Fundamentos, 1-3
Teoria das três idades, 2
B
Biblioteca, 40, 43-45, 52, 60, 61, 63- 64, 75
Ver também arquivo
Biblioteca Escolar, 41
Biblioteca Especializada, 52
Biblioteca Nacional, 60
Biblioteca Pública, 63-64
Biblioteca Universitária, 75
Conceitos, 40
Estrutura, 43-44
Planejamento, 40
Serviços, 44-45
Bibliotecário, 41, 44-45, 52, 75, 85
Bibliotecário escolar, 41
Bibliotecário especializado, 52
Bibliotecário universitário, 75
Legislação, 85
Mercado de trabalho, 44-45
Biblioteconomia, 40-43, 51, 63
[...]
Z
Zona urbana, 41, 43
Zona rural, 41, 43

Fonte: Produção do próprio autor (2015).

Outro tipo de índice importante é o **índice onomástico**, cujas entradas são exclusivamente formadas por nomes de autores. Esse tipo de índice possui arranjo alfabético por sobrenome dos autores. Alguns índices são híbridos, isto é, apresentam entradas para temáticas e sobrenomes de autores.

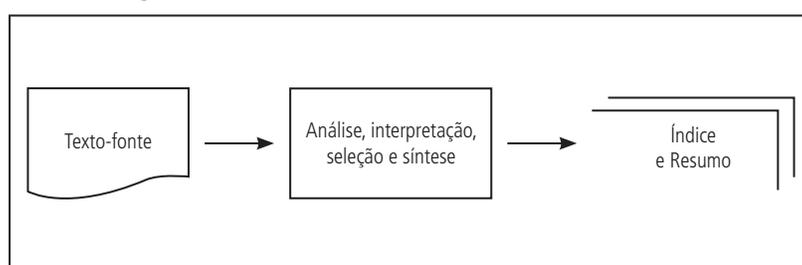
Existem também índices que se constituem em verdadeiros documentos referenciais. Essas são publicações analíticas elaboradas por serviços especializados de indexação e resumo, que assumem o formato de resumos (*abstracts*). Como exemplos, destacam-se o *Chemical Abstracts* (CA) e o *Library and Information Abstracts* (LISA), por se tratar de documentos mais conhecidos nas áreas de Química e Ciência da Informação.

No universo digital, também são usadas informações documentárias denominadas de ontologias, que têm como objetivo organizar os conteúdos disponíveis no meio eletrônico, visando à sua posterior recuperação. Esse tipo de produto é também bastante completo, envolvendo atividades de catalogação, classificação e indexação, entre outras (DIAS; NAVES, 2007).

O **resumo** é outro tipo de produto que resulta dos processos de análise e síntese. Do ponto de vista material, apresenta o mesmo conceito de índice, diferenciando-se deste apenas do ponto de vista formal, na medida em que se refere a um texto dissertativo, todo coeso, com início, meio e fim, conforme Figura 16.

A indexação de assuntos e a redação de resumos são atividades intimamente relacionadas, pois ambas implicam a preparação de uma **representação** do conteúdo temático dos documentos. O resumidor redige uma descrição narrativa ou síntese do documento, e o indexador descreve seu conteúdo ao empregar um ou vários termos de indexação, comumente selecionado de algum tipo de vocabulário controlado (LANCASTER, 2004, p. 6, grifo do autor).

Figura 16 – Índice e resumo: materialidade e formalidade

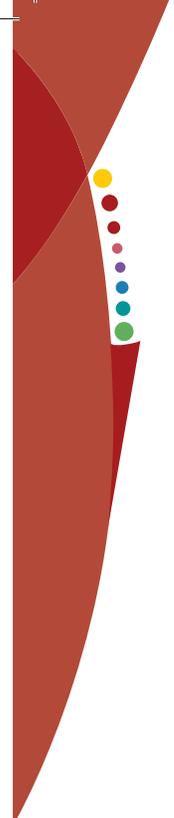


Fonte: Produção do próprio autor (2015).

A partir de Lancaster (2004), considera-se que o resumo é a apresentação concisa, porém exata, do conteúdo de um documento. Não se confunde com mero extrato, que corresponde à extração e à compilação de trechos de um documento, não representando, portanto, um novo texto.

A compreensão do resumo pode ser realizada a partir das características do índice porque, do ponto de vista material e procedimental, indexação e resumos se assemelham bastante. Trata-se, conforme Figura 16, das operações consecutivas de análise, interpretação, seleção e síntese





dos conteúdos dos documentos. A diferença, contudo, se encontra na formalidade de apresentação, uma vez que a indexação usa termos indexadores ou palavras-chave e o resumo, texto dissertativo.

De acordo com a NBR 6028/ABNT (2003), há três tipos de resumos, a saber, **indicativo**, **informativo** e **crítico**. Esses tipos de resumos se diferenciam conforme a finalidade e a forma de representação do texto-fonte.

O **resumo indicativo** é composto por uma breve descrição dos pontos principais do texto-fonte, sem expressar dados qualitativos e/ou quantitativos do conteúdo sintetizado. Em função disso, esse tipo de resumo, de modo geral, não dispensa a consulta ao objeto de análise e síntese. Você constatará que essa categoria de resumo é frequentemente adotada na elaboração de artigos científicos, comunicações científicas e serviços de alerta, e são comumente elaborado pelos próprios autores dessas publicações.

O **resumo informativo** é composto por uma explicação sucinta sobre todos os pontos principais do texto-fonte, quais sejam as finalidades, a metodologia adotada no trabalho, os resultados alcançados e as conclusões ou considerações finais. Dada a representação de todas as partes, o acesso a este tipo de resumo, muitas vezes, dispensa a consulta ao texto-fonte. Você observará que esta categoria de resumo é frequentemente usada na elaboração de monografias (*stricto sensu*), dissertações, teses, relatórios técnico-científicos e base de dados referenciais. Nesse domínio, o resumo informativo tem a mesma autoria dessas publicações.

Lancaster (2004) esclarece que os resumos informativos são comumente mais extensos que os indicativos e mais difíceis de serem redigidos, considerando, sobretudo, o tipo de texto-fonte. Em se tratando de um estudo experimental, torna-se relativamente fácil analisar, identificar e sintetizar os elementos que compõem um resumo informativo. Por outro lado, em se tratando de um estudo teórico ou opinativo, por vezes, torna-se impossível contemplar alguns elementos que caracterizam esse tipo de resumo. Em decorrência disso, os resumos informativos são mais frequentes nas ciências exatas e tecnológicas do que nas ciências sociais e humanas.

O **resumo crítico** apresenta uma análise ponderada do conteúdo informativo do texto-fonte. Em função de esse tipo de síntese ter como finalidade expressar um julgamento do conteúdo informativo do documento em análise, com autoria distinta, deve ser redigido por um especialista na área. Essa categoria de resumo é genericamente também chamada de resenha, mas, quando analisa uma edição de um documento, entre várias edições, denomina-se *recensão*.

Além desta classificação, *Lancaster* (2004) apresenta outros tipos de resumo, a saber: **literatura concisa**, **miniresumo**, **resumo estruturado**, **resumo modular** e **resumo telegráfico**.

A **literatura concisa** é um tipo de resumo cujo enunciado altamente sucinto tem como finalidade sintetizar o ponto principal do conteúdo de um documento. Há ainda uma variação desse tipo de resumo, a conclusão ultraconcisa, que corresponde à principal conclusão de uma pesquisa. Não se trata efetivamente de um resumo, porém, acompanhado da referência do documento a que se refere, pode desempenhar importante papel na identificação deste e de conteúdos específicos. Segundo *Lancaster* (2004), este tipo de resumo é estruturado em torno de três questões, a saber, “o quê?”, “como?” e “por quê?”, conforme Figura 17.

Figura 17 – Literatura concisa

A Ciência da Informação é um campo interdisciplinar porque tem por base a complexidade do objeto de estudo e a formação dos pesquisadores que atuam no campo informacional.

Fonte: Produção do próprio autor (2015).

O **miniresumo** apresenta uma nomenclatura bastante imprecisa, na medida em que apenas se refere a um resumo curto no estilo de mensagem telegráfica, como pode ser observado na Figura 18. Tomando como referência Lunin (1967), Lancaster (2004) esclarece que a expressão se refere a um resumo altamente estruturado usado na realização de buscas automáticas. Em outros termos, refere-se a uma espécie de cruzamento entre um resumo e uma entrada de índice, que resulta em uma sequência de termos muito próxima à estrutura de uma frase. Assim, o enunciado “a indexação tem como finalidade representar o conteúdo de um documento”, por exemplo, seria resumida da seguinte forma.

Figura 18 – Miniresumo

INDEXAÇÃO/REPRESENTAÇÃO/CONTEÚDO/DOCUMENTO

Fonte: Produção do próprio autor (2015).

O **resumo estruturado** se aproxima bastante de um índice, assemelhando-se a um gabarito. Na elaboração deste, devem-se procurar especificamente os itens listados. Embora seja um tipo de resumo bastante simples, para algumas finalidades este tipo de resumo é preferível ao formato narrativo, como, por exemplo, na compilação de manuais. A Figura 19 se refere ao gabarito estruturado de um documento cujo conteúdo central é “tipos de representação”. A elaboração do resumo é realizada pelo registro de “valores” apropriados no gabarito.

Figura 19 – Resumo estruturado

TIPO DE REPRESENTAÇÃO	INSTRUMENTOS	PROCESSOS	PRODUTOS

Fonte: Adaptado de LANCASTER (2004, p. 106).

O **resumo modular** é o mais completo de todos os tipos apresentados por Lancaster (2004), compreendendo inclusive outros tipos de resumo. Os resumos modulares têm a finalidade de fornecer uma descrição completa de documentos correntes, evitando, assim, desperdícios de esforço intelectual, na medida em que pode ser utilizado por diferentes serviços de resumo. Este tipo de resumo é formado por cinco partes: uma citação,



uma anotação, um resumo indicativo, um resumo informativo e um resumo crítico, conforme Figura 20. A citação que se encontra nos exemplos de resumos modulares, no contexto brasileiro, se refere à referência, que deve ser elaborada segundo a NBR 6023/ABNT (2018). A anotação corresponde a um enunciado geral que sintetiza o conteúdo informativo do texto-fonte.

Figura 20 – Resumo modular

Referência

SOUZA, E. D. Configurações do campo da Ciência da Informação: pluralismo epistemológico e descentração interdisciplinar. **Tendências da pesquisa brasileira em Ciência da Informação**, v. 5, n. 1, 2012. Disponível em: <<http://inseer.ibict.br/ancib/index.php/tpbci/article/view/63/104>>. Acesso em: 15 set. 2015.

Anotação

A epistemologia interdisciplinar, na Ciência da Informação, promove o pluralismo epistemológico e a descentração interdisciplinar, que resultam em flexibilidade e alargamento do campo científico.

Indicativo

*As configurações de um campo científico decorrem das condições de produção em que se inscreve e das práticas científicas nele desenvolvidas. No campo da Ciência da Informação, a epistemologia interdisciplinar se destaca, uma vez que se apresenta como um de seus principais fundamentos. Assim, busca-se compreender os reflexos dessa prática científica na consolidação do seu campo disciplinar. Para tanto, define-se como referencial teórico-metodológico a *Análise do Discurso* de Michel Pêcheux e como fonte material de análise 83 (oitenta e três) artigos publicados nos principais periódicos brasileiros, no período de 1990 a 2008. Considera-se que a epistemologia interdisciplinar promove o pluralismo epistemológico e a descentração interdisciplinar, que resulta em flexibilidade e alargamento do campo científico.*

Informativo

*A Ciência da Informação tem como condições de produção o pluralismo epistemológico e o estabelecimento do campo interdisciplinar, que decorrem da complexidade do seu objeto de estudo e da formação de seus pesquisadores. A epistemologia interdisciplinar se destaca, na medida em que se apresenta como um de seus principais fundamentos. Assim, busca-se compreender os reflexos dessa prática científica na consolidação do seu campo disciplinar, especificamente, no que se refere ao pluralismo epistemológico e à prática de descentração interdisciplinar. Para tanto, define-se como referencial teórico-metodológico a *Análise do Discurso* de Michel Pêcheux e como fonte material de análise 83 (oitenta e três) artigos publicados nos principais periódicos brasileiros, no período de 1990 a 2008. Procurou, a partir da análise do discurso, o estabelecimento de relações entre o já dito em outro lugar e momento (interdiscurso) e o que estava sendo dito (intradiscurso). O pluralismo epistemológico intimida a identidade da Ciência da Informação porque é colocado em uma abordagem a-histórica naturalizante e em uma filosofia reflexivo-idealista do fenômeno informacional, não possibilitando a emergência das condições disciplinares em que essa interdisciplinaridade acontece. Esse pluralismo contribui, por um lado, com a flutuação, e, por outro, com a dispersão, que dificultam o processo de integração disciplinar.*

A prática da descentração é resultado da necessidade de estudar problemas complexos que não são possíveis de ser abordados no interior de uma disciplina. Essas práticas de descentração não têm uma disciplina como ponto de partida e também não têm um ponto de chegada. Elas se desenvolvem, por conseguinte, em uma constante flutuação teórico-metodológica. Conclui que o pluralismo epistemológico promove a flexibilização e o alargamento do campo disciplinar, buscando acomodar os diversos pontos de vista que resultam no desenho do campo sustentado pela epistemologia interdisciplinar.

Crítico

A Ciência da Informação é apontada como um campo interdisciplinar por natureza, em função do pluralismo que a caracteriza. Essa constatação vem sendo feita por diversos pesquisadores, desde a realização dos primeiros estudos na segunda metade do século passado. Esses estudos adotam, em grande medida, uma abordagem empírico-positivista, uma vez que faz uso de processos quantitativos visando a identificar as áreas de cooperação disciplinar. Para tanto, concentram nos estudos bibliométricos acerca das principais temáticas presentes na literatura da Ciência da Informação. O presente trabalho, contudo, adota uma abordagem crítica que tem como referências as condições de produção e os aspectos qualitativos da produção científica da Ciência da Informação. Nesse sentido, procura, a partir da abordagem materialista do discurso, compreender as implicações da prática interdisciplinar na consolidação do campo da Ciência da Informação, considerando as relações entre o já dito (interdiscurso) e o que estava sendo dito (intradiscurso). As análises e discussões têm como base material 83 (oitenta e três) artigos científicos publicados nos principais periódicos brasileiros. A partir das condições amplas e específicas da produção analisada, constata-se que o pluralismo epistemológico dificulta a identidade disciplinar da Ciência da Informação porque tem por base uma abordagem a-histórica naturalizante e uma filosofia reflexivo-idealista do fenômeno informacional, impossibilitando a emergência das condições disciplinares em que essa interdisciplinaridade acontece. Esse pluralismo contribui, a um só tempo, com a flutuação e a dispersão, que dificultam o processo de integração disciplinar. A prática da descentração é constatada, por conseguinte, como resultado da necessidade de estudar problemas complexos que não são possíveis de ser abordados no interior de uma única disciplina. Essas práticas de descentração se caracterizam pelo movimento errante, resultando em constante flutuação teórico-metodológica. Conclui que o pluralismo epistemológico promove a flexibilização e o alargamento do campo disciplinar, buscando acomodar os diversos pontos de vista que resultam no desenho do campo sustentado pela epistemologia interdisciplinar. Considerando o estágio em que se encontra o campo científico da Ciência da Informação, o estudo considera as práticas interdisciplinares como obstáculos epistemológicos à consolidação do seu campo disciplinar.

Fonte: Adaptado de LANCASTER (2004, p. 108).

O **resumo telegráfico**, por fim, também se refere a uma imprecisão conceitual, pois não representa o conteúdo de um documento na estrutura de enunciados telegráficos, como ocorre com os minirresumos. Trata-se de uma cadeia de termos altamente sucinta e desprovida de sintaxe, que foi usada para designar um componente do primeiro sistema de recuperação automatizado. Este tipo de resumo era elaborado de acordo com um conjunto de regras preestabelecidas a fim de excluir variações e complexidades da estrutura da frase em língua inglesa. Para tanto, os termos adotados eram codificados por intermédio de um dicionário de código semântico, atribuindo-se a cada radical quatro dígitos formados por três caracteres e um espaço para interpolação de um quarto caractere, conforme a Figura 21.



Figura 21 – Resumo telegráfico

C-TL Catalyst [Catalisador]
C-TR Container [Recipiente]
C-TT Cutting and drilling [Corte e perfuração]
D-DD Damage [Dano]
D-FL Deflection [Desvio]

Fonte: LANCASTER (2004, p. 192).

Conhecidos os diversos tipos de resumo, você pode observar que as atividades de indexação e redação de resumos estão intimamente relacionadas, desde os procedimentos adotados na representação da informação às finalidades vinculadas aos processos de recuperação da informação. Por um lado, é imprescindível compreender que não há sobreposições, mas complementaridades entre essas duas modalidades de representação temática da informação. É também importante considerar que, na literatura técnico-científica brasileira, os tipos de resumos adotados são aqueles que são objeto da NBR 6028/ABNT (2003), isto é, indicativo, informativo e crítico, que apresentam bastante diferença em relação aos cabeçalhos de assuntos, descritores ou termos elaborados no processo de indexação. Entretanto, em periódicos na área de saúde indexados em bases de dados como, por exemplo, a *SciELOBrazil*, verifica-se o uso recorrente de resumo estruturado.

Segue, na Figura 22, um esquema que lhe orientará sobre os principais componentes e os procedimentos que devem ser considerados na redação de resumos.

Figura 22 – Esquema de elaboração de resumos

1. Tipo de resumo, conforme a finalidade.
2. Extensão do resumo, conforme o seu tipo e a NBR 6028/ABNT (2003).
3. Terminologia especializada ou técnica do texto-fonte.
4. Conteúdo:
 - a) Tópico frasal, enunciando (resumo indicativo) ou contextualizando (resumos informativo e crítico);
 - b) Objetivo geral (todos os tipos de resumo) e objetivos específicos (resumos informativo e crítico) da pesquisa;
 - c) Procedimentos metodológicos;
 - d) Resultados da pesquisa;
 - e) Conclusões e/ou considerações finais.
5. Elabore orações em lugar de períodos, frases em lugar de orações, palavras em lugar de frases, quando possível.
6. Não usar:
 - a) Abreviaturas e/ou siglas;
 - b) Símbolos de caráter não convencional;
 - c) Equações ou nota de rodapé;

- d) Recuos;
- e) Dados de catalogação descritiva.
7. Verbo na voz ativa.
8. Verbo na terceira pessoa do singular.
9. Precedido de referência, conforme a NBR 6023/ABNT (2018), sempre que não for elemento pré-textual do próprio documento (artigos, monografias e relatórios).
10. Sucedido por palavras-chave, que são antecedidas da expressão “palavras-chave”; todas separadas entre si por um ponto e também finalizadas por um ponto.
11. Classificação de sigilo, se houver algum controle de acesso.
12. Revise-o, segundo os padrões linguísticos vigentes e as normas, NBR 6023/ABNT (2018) e NBR 6028/ABNT (2003).

Fonte: Adaptado de LANCASTER (2004, p. 115).

Ainda sobre a elaboração de resumos, *Lancaster* (2004) acrescenta que a inclusão de elementos na elaboração de um resumo depende do tipo de publicação pretendida, considerando inclusive as áreas de conhecimento e o tipo de texto-fonte, que pode abordar de questões bastante concretas e/ou práticas a altamente abstratas e/ou teóricas, conforme discutido anteriormente. Assim, em que pese o apontamento dos principais componentes na Figura 22, há diversas possibilidades de elaboração de resumos. Nesse domínio, o autor aponta uma tendência mais recente à elaboração de **resumos altamente formatados**, destacando as partes componentes com a adoção de parágrafos e a atribuição de entretítulos (introdução, método, resultados e conclusão), conforme a Figura 23. Esse tipo de resumo, diferentemente da classificação apresentada por este autor (ver Figura 19), vem sendo também denominado de “resumo estruturado” e adotado, principalmente, nos periódicos científicos da área de medicina.

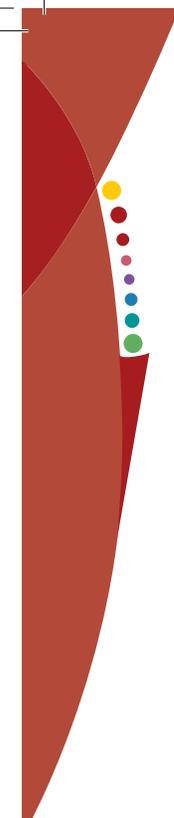
Figura 23 – Resumo altamente formatado

ROSAS, P. et al. Adequação dos descritores em ciências da saúde para a indexação de dissertações acadêmicas, na área de doenças respiratórias. **Jornal Brasileiro de Pneumologia**, Brasília, v. 25, n. 6, p. 309-312, nov./dez. 1999. Disponível em: <<http://www.scielo.br/pdf/jpneu/v25n6/v25n6a03.pdf>>. Acesso em: 10 nov. 2015.

Introdução: O processo de indexação consiste em duas fases: primeira, identificar e representar o conteúdo intelectual de um documento; segunda, traduzir a análise do assunto para uma linguagem específica, utilizando descritores. Na biblioteca do Instituto de Doenças do Tórax da UFRJ, a indexação era realizada por meio de um vocabulário controlado – DeCS (Descritores em Ciências da Saúde). O objetivo deste trabalho foi verificar se são adequados, para indexação, os descritores empregados pelos autores de teses e dissertações de mestrado, na área de doenças respiratórias. **Método:** Os descritores de 29 teses e dissertações de mestrado em Tisiologia e Pneumologia, de 1990 a 1996, foram estudados. **Resultados:** A coleta de dados evidenciou que 29 autores empregaram 101 descritores (3,48 descritores/autor). Dos 101 descritores, 47 (47%) eram adequados. **Conclusão:** Quarenta e sete por cento dos descritores empregados, nas teses e dissertações do Curso de Mestrado em Tisiologia e Pneumologia do IDT-UFRJ (1990 a 1996), são adequados segundo o DeCS.

Fonte: Jornal Brasileiro de Pneumologia (1999).





Alguns periódicos científicos brasileiros adotam essa pré-formatação dos resumos como, por exemplo, o *Jornal Brasileiro de Pneumologia*, publicado pela *Sociedade Brasileira de Pneumologia e Tisiologia*. A redação deste tipo de resumo pode apresentar algumas facilidades para os consultentes, principalmente nas buscas informatizadas e na identificação de artigos vinculados às respectivas práticas profissionais. Por outro lado, os pontos críticos deste tipo de resumo, apontados por *Lancaster* (2004), a partir de vários estudos, devem-se ao aumento da extensão, em média 21%, e às pretensões por vezes exageradas dos resumos. O fato é que, na prática, alguns autores, no momento da redação do resumo, podem sentir-se obrigados a incluir elementos além da presença efetiva destes no respectivo texto-fonte, na medida em que procuram contemplar a estrutura exigida pelas normas editoriais da publicação.

2.4 PRÁTICA DA ANÁLISE DA INFORMAÇÃO

Do ponto de vista prático, a análise da informação se refere ao conjunto de procedimentos realizados com a finalidade de identificar o conteúdo de um documento. De forma bastante clara, trata-se de procurar responder à seguinte pergunta: qual é o conteúdo principal deste documento?

Não há uma metodologia única que aponte um conjunto de procedimentos que devem ser adotados na análise da informação, em função da diversidade de elementos que a condicionam e da complexidade de algumas atividades que a compõem. Alguns procedimentos mais comuns podem, contudo, ser aqui abordados.

Primeiro, deve-se ter em mente que:

A representação documentária é obtida por meio de um processo que se inicia pela análise do texto, com o objetivo de identificar conteúdos pertinentes em função das finalidades do sistema – e da representação desses conteúdos – numa forma sintética, padronizada e unívoca (CINTRA *et al.*, 2002, p. 38-39)

Conforme abordado anteriormente, é imprescindível constar que o conteúdo informacional pode ser expresso de diferentes formas, tais como texto, som, imagem e combinações diversas. Essas formas de expressão compõem a base material de trabalho do analista da informação.

A partir daí, seguindo *Lancaster* (2004), na condição de analista da informação, você deve formular, pelo menos, três perguntas e procurar respondê-las:

- a) Qual o conteúdo informativo deste documento?
- b) Por que este foi incorporado ao acervo da biblioteca?
- c) Quais de seus aspectos serão de interesse para os usuários desta biblioteca?

A análise da informação tem, portanto, três universos de referência, o conteúdo da fonte, a relação deste com a finalidade da instituição bibliotecária e com o acervo já existente, e, principalmente, o interesse da comunidade usuária. Em que pesem os elementos subjetivos (analista da informação e usuários) presentes nesses questionamentos e nas suas respectivas respostas, considerá-los inicialmente na prática da análise da informação significa uma aproximação à maior qualidade e à melhor eficiência nos processos de tratamento, organização e recuperação da informação.

No terceiro momento, o analista deve procurar compreender a base estrutural, visando ao conhecimento prévio do texto sobre o qual irá realizar a análise:

- a) **superestrutura:** esquema de organização textual, conforme a sua tipologia;
- b) **macroestrutura:** tópicos representativos que expressam a organização semântica e sintática do texto e
- c) **microestrutura:** estrutura superficial do texto.

Essa primeira aproximação da base material das análises deve considerar a política de indexação e elaboração de resumos que define as medidas de extensão da representação da informação quanto à abrangência semântica (geral ou específica) e à cobertura do conteúdo do documento (seletiva ou exaustiva). A consideração prática destas medidas é indispensável porque estas irão interferir diretamente na eficiência da recuperação da informação.

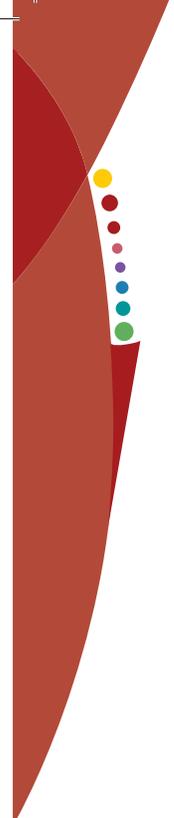
Essas decisões estão relacionadas a um conjunto de princípios, variáveis e estratégias que vão desde a identificação do usuário à realização de busca em bases de dados. Com efeito, faz-se necessária a adoção de uma política de indexação que busque “Definir essas variáveis, estabelecer princípios e critérios que servirão de guias na tomada de decisões para otimização do serviço, racionalização dos processos e consistência das operações neles envolvidas [...]” (CARNEIRO, 1985, p. 222).

Uma política de indexação deve, conforme *Carneiro* (1985), contemplar os seguintes elementos: cobertura de assuntos; seleção e aquisição de documentos, processo de indexação (linguagem, medidas de extensão e coeficientes de recuperabilidade); estratégias de busca; tempo de resposta do sistema; formas de saída e avaliação do sistema. Essas questões deverão ser tratadas com maior profundidade em disciplinas que se dedicam especificamente aos processos, instrumentos e produtos da representação temática da informação.

Outra questão que definirá a prática analítica diz respeito ao tipo de representação que será realizada, temática (classificação, indexação e/ou resumo) e descritiva (catalogação e/ou referência). Dependendo do tipo de representação, o analista se dedicará a informações documentárias específicas.

Ademais, deve-se considerar que, em função da quantidade de material para o tratamento, a análise deve ser feita a partir da leitura documentária, isto é, não deve realizar a leitura integral do documento, salvo em se tratando da elaboração de índice que integra o texto-fonte como, por exemplo, índice alfabético de livro.





A partir daí, o analista realizará a operacionalização da análise informacional em três etapas:

- a) compreensão das informações documentárias descritivas e/ou temáticas;
- b) identificação das informações documentárias descritiva e/ou temáticas e
- c) seleção das informações documentárias descritivas e/ou temáticas.

Na primeira etapa, o analista deve examinar o documento, conforme o tipo de representação. Assim, pode se ater às informações descritivas, nos termos do AACR2 e/ou da NBR 6023/ABNT, e/ou às informações temáticas que servirão de base para os processos de classificação, indexação e/ou resumo. Essas informações podem ser obtidas a partir de título, subtítulo, sumário, apresentação, gráficos, tabelas e outras partes do texto que mereçam atenção.

Na segunda etapa, o analista deve, naquele conjunto de informações documentárias, detectar as categorias que melhor representam, extrinsecamente e intrinsecamente, o documento. Do ponto de vista descritivo, essa é uma atividade relativamente fácil porque tem como referência informações documentárias objetivas. No domínio temático, a tarefa se torna mais complexa porque exigirá do analista o estabelecimento de relações sistemáticas entre as categorias de assuntos ou os conceitos que poderão ser usados na representação, considerando a garantia literária e a garantia do usuário.

Nessa etapa, o analista pode usar instrumentos auxiliares, tais como sistemas de classificação, listas de cabeçalhos de assunto e tesouros, visando à melhor compreensão do domínio conceitual e das relações entre os conceitos que constituem os diferentes assuntos.

Na terceira e última etapa, o analista deve definir as informações documentárias que representam o documento. Como destacado acima, os elementos descritivos são relativamente fáceis de serem selecionados porque têm por base critérios objetivos, isto é, o que está disposto nos padrões normativos, AACR2 e/ou NBR 6023/ABNT. No domínio temático, torna-se mais difícil em função da presença da subjetividade no processo seletivo, porém, o analista deve, na medida do possível, primar pela imparcialidade com a adoção sistemática dos principais universos de referência, a garantia literária e a garantia do usuário.

Nessa etapa, o analista deve assumir uma postura vigilante porque ele tende a selecionar conteúdos e conceitos de seu domínio prévio e, até mesmo mais familiares, inclusive em função da adoção de estratégias cognitivas. Trata-se, diga-se de passagem, de prestar atenção à identificação de novos conceitos que podem ampliar o domínio do sistema de recuperação da informação.

Além disso, *Fujita* (2004) esclarece que a leitura é uma prática social realizada por intermédio de complexos processos de interação entre o leitor e o texto, em um contexto específico. Nessa perspectiva, a leitura documentária é uma prática profissional do indexador composta por três elementos, a saber, o **leitor**, o **texto** e o **contexto**. Em função das particularidades que assumem em cada ato de leitura, esses componentes transformam-se em variáveis. A variável leitor compreende o conhecimento prévio do leitor-indexador sobre o assunto e as estratégias de

leitura que este utiliza. A variável texto corresponde ao material a ser lido e abarca os seguintes aspectos: a intenção do autor, a estrutura textual, a escrita e o conteúdo temático. A variável contexto corresponde aos elementos extratextuais que podem influenciar na compreensão da leitura, tais como intenção da leitura, interesse na leitura, intervenções de terceiros, tempo e barulho. Esse contexto compreende, portanto, atributos físicos, psicológicos e sociais, entre outros. A autora esclarece ainda que, quanto maior for a interação entre essas variáveis na leitura, melhor será o nível de compreensão do texto.

Finalmente, é importante que, antes de prosseguir para a segunda etapa da representação, a tradução, o analista revise todo o processo analítico, objetivando a detecção de possíveis inconsistências, sobretudo em relação ao acervo existente e às linguagens documentárias adotadas no sistema de recuperação da informação.

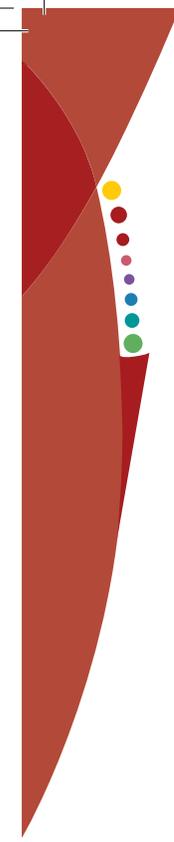


RESUMO

O sistema de recuperação da informação procura estabelecer a comunicação entre dois universos povoados de singularidades, **universo da informação** e **universo do usuário**, que se caracterizam pelas dispersões semânticas e sintáticas. A representação temática da informação pode ser realizada com o uso da **linguagem natural** ou de uma **linguagem artificial**. A primeira corresponde à linguagem adotada pelos autores na composição dos conteúdos informacionais e pelos usuários na elaboração de suas demandas de informação. A segunda, a um conjunto de termos e/ou símbolos destinados à representação dos conteúdos dos documentos em sistemas de recuperação da informação, por isso são denominadas de linguagem documentária. Esta linguagem é composta por dois componentes básicos, a saber, **vocabulário** e **sintaxe**, e tem como finalidade realizar o controle de vocabulário da linguagem natural, permitindo melhor comunicação entre usuário e sistema. No domínio do vocabulário, há três categorias de relações semânticas, **relações de equivalência** (sinônimos, antônimos, grafias diferentes, abreviaturas e siglas, e traduções), **relações hierárquicas** (gênero e espécie, e todo e parte) e **relações afinitivas** ou **associativas** (causa e efeito, concorrência, coordenação, genética, instrumental, similaridade de processos e produtos, e substância). No domínio da sintaxe, o controle é realizado na composição da estrutura frásica (cabecinhos de assuntos, notações e descritores) de conceitos compostos e complexos, por intermédio da coordenação de conceitos simples. As linguagens documentárias são classificadas, segundo o momento em que os termos ou símbolos são coordenados, em **linguagens pré-coordenadas** e **linguagens pós-coordenadas**. Nas primeiras, a coordenação é realizada no momento da representação da informação e tem como principais exemplos as **listas de cabecinhos de assunto** (*Library of Congress of Subject Headings* e *Sears List of Subject Headings*) e os **sistemas de classificações bibliográficas** (*Classificação Decimal de Dewey* e *Classificação Decimal Universal*). Nas últimas, a coordenação é realizada no momento da recuperação da informação e tem como principal exemplo o **tesauro**. As linguagens documentárias expres-

- I. A indexação por extração, conhecida também como indexação por conceitos, é realizada usando a linguagem natural.
- II. A indexação por atribuição, conhecida também por indexação por palavras, é realizada usando a linguagem artificial.
- III. A indexação por atribuição, conhecida também como indexação por conceito, é realizada usando a linguagem artificial.
- IV. A indexação por extração, conhecida também como indexação por palavras, é realizada usando a linguagem natural.
- Estão corretas:
- a) I e II, apenas.
- b) II e III, apenas.
- c) III e IV, apenas.
- d) II e IV, apenas.
- e) I, II, III e IV.
3. As linguagens de informação, assim como as linguagens naturais, são formadas por dois componentes básicos. Marque a alternativa correta.
- a) O vocabulário e a sintaxe são os elementos básicos e designam, respectivamente, os controles sintáticos e semânticos na representação da informação.
- b) A semântica e a sintaxe são os elementos básicos e definem, respectivamente, a estrutura frásica e o significado dos termos indexadores.
- c) O vocabulário e a sintaxe são os elementos básicos e têm por objeto, respectivamente, as relações semânticas e a ordem de citação.
- d) O vocabulário e a semântica são os elementos básicos e têm por objeto, respectivamente, o controle sintático e a ordem de citação.
- e) A semântica e a sintaxe são os elementos básicos e expressam, respectivamente, a estrutura frásica e a ordem de citação nos cabeçalhos de assunto.
4. O controle de vocabulário na representação da informação tem por base dois tipos de relações, a saber, coordenadas e subordinadas. Assinale a opção correta.
- a) As relações coordenadas são realizadas entre elementos do mesmo domínio e de níveis hierárquicos diferentes.
- b) As relações coordenadas e subordinadas são realizadas entre elementos de domínios e níveis hierárquicos diferentes.
- c) As relações subordinadas são realizadas entre elementos de domínios e níveis hierárquicos diferentes.
- d) As relações coordenadas são realizadas entre elementos de mesmos domínios e níveis hierárquicos.
- e) As relações coordenadas e subordinadas são realizadas entre elementos de mesmos domínios e níveis hierárquicos diferentes.



- 
5. A definição das relações existentes entre termos indexadores é fundamental ao controle de vocabulário nos sistemas de informação. Analise as afirmativas e marque a opção correta.
- I. As relações hierárquicas são tipos de relações restritivas, que ofertam termos alternativos na recuperação da informação.
 - II. As relações associativas são espécies de relações ampliativas, que ofertam termos alternativos para representação da informação.
 - III. As relações de equivalência são espécies de relações restritivas, que ofertam termos alternativos na recuperação da informação.
 - IV. As relações afinativas são tipos de relações ampliativas, que ofertam termos alternativos na recuperação da informação.

Estão corretas:

- a) I e II, apenas.
 - b) I e III, apenas.
 - c) II e IV, apenas.
 - d) III e IV, apenas.
 - e) I, II, III e IV.
6. A coordenação é uma estratégia usada para especificar a representação da informação. Analise as alternativas e aponte a correta.
- a) Na pré-coordenação, o usuário do sistema especifica o conteúdo do documento.
 - b) A pós-coordenação é uma eficiente estratégia de exclusão de falsas coordenações.
 - c) A pré-coordenação tem como princípio básico a adoção de conceitos simples.
 - d) A pós-coordenação é uma eficiente estratégia de fixação das relações entre dois conceitos.
 - e) A pré-coordenação realiza a ordem de citação e promove a precisão na recuperação da informação.
7. As linguagens documentárias são sistemas de controle de linguagem que exercem diferentes papéis na organização e recuperação da informação. Analise os itens seguintes:
- I. As classificações bibliográficas são linguagens que objetivam estabelecer a relação entre os documentos e indicar as respectivas localizações no acervo.
 - II. O tesouro é um instrumento de controle que representa as informações descritivas fazendo uso de conceitos simples.
 - III. A lista de cabeçalhos de assunto é um instrumento de controle que representa as informações descritivas fazendo uso de conceitos compostos.
 - IV. As classificações bibliográficas e as listas de cabeçalhos de assunto são linguagens que prefixam as estruturas sintáticas das representações.

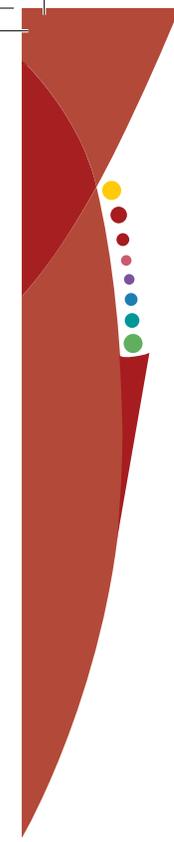
No conjunto de assertivas acima, estão corretas:

- a) I e II, apenas.
 - b) I e III, apenas.
 - c) II e III, apenas.
 - d) I e IV, apenas.
 - e) I, II, III e IV.
8. A classificação, a indexação e a elaboração de resumos são três tipos de representação que guardam semelhanças e diferenças. Analise as afirmativas abaixo:
- I. A classificação e a elaboração de resumos são processos de representação descritiva, que se diferenciam porque a primeira tem como produto a notação e a segunda, um texto dissertativo.
 - II. A indexação e a elaboração de resumos são processos de representação temática, que se diferenciam nos produtos, respectivamente, termos indexadores e resumos.
 - III. A classificação e a indexação são processos de representação temática, que se diferenciam nos sistemas de notação, respectivamente, sistemático e alfabético.
 - IV. A classificação, a indexação e a elaboração de resumos são processos semelhantes de representação, sendo a classificação e a indexação formas de representação temática e a elaboração de resumo uma forma descritiva.

As assertivas corretas são:

- a) I e III, apenas.
 - b) II e III, apenas.
 - c) II e IV, apenas.
 - d) I, II, III e IV.
 - e) III e IV, apenas.
9. O resumo corresponde a uma representação sucinta do conteúdo informativo dos documentos. Marque a opção correta.
- a) O resumo indicativo é uma descrição precisa do conteúdo do documento que, em determinadas situações, substitui o texto-fonte.
 - b) O resumo crítico, que também pode ser chamado de resenha ou recensão, deve preferencialmente ser elaborado por um especialista.
 - c) O resumo informativo é composto por uma breve descrição do conteúdo do documento e, em geral, não dispensa a leitura do texto-fonte.
 - d) O resumo telegráfico é desprovido de sintaxe, por conseguinte, se apresenta no formato de mensagem telegráfica.
 - e) A literatura concisa é tipo complexo de resumo da informação composta por referência, anotação, resumo indicativo, resumo informativo e resumo crítico.
10. A análise da informação é uma atividade complexa do processo de representação da informação, que exige do analista





conhecimento prévio do domínio conceitual e atenção nas diferentes atividades. Leia as assertivas abaixo:

- I. A identificação das informações documentárias temáticas é realizada a partir do exame dos elementos extrínsecos do documento.
- II. A seleção das informações documentárias descritivas é operacionalizada com base em critérios objetivos porque segue padrões normativos.
- III. A compreensão das informações documentárias temáticas é realizada a partir do exame dos elementos intrínsecos do documento e é caracterizada pela subjetividade.
- IV. A seleção das informações documentárias descritivas exige a análise sistemática das relações entre categorias de assuntos ou conceitos.

Quais estão corretas?

- a) I e III, apenas.
 - b) I e IV, apenas
 - c) II e III, apenas.
 - d) II e IV, apenas.
 - e) I, II, III e IV.
-

UNIDADE 3

INTERDISCIPLINARIDADE NA ANÁLISE DA INFORMAÇÃO





O desenvolvimento das sociedades tem proporcionado a ampliação e o aprofundamento dos diferentes tipos de conhecimento – mítico, teológico, filosófico, científico, técnico, estético, entre outros –, que resultaram no aumento dos diversos estoques de documentos cuja disponibilização depende de infraestruturas planejadas.

Atualmente a análise da informação é realizada em busca do estabelecimento de redes de representação dos conteúdos informacionais, que visam a construir representações inteligíveis do conhecimento e estabelecer relações possíveis entre estratos de informações que compõem os sistemas de recuperação da informação. O conhecimento interessa à Biblioteconomia e às áreas afins, principalmente, no seu caráter representacional, na medida em que é a partir do conjunto de representações que se estruturam e se organizam os diversos sistemas de recuperação da informação, tais como bibliotecas, centros de documentação e centros de informação.

A análise da informação se localiza no campo de estudo da memória documentária, que, segundo *Dodebei* (2002), tem como fundamentos específicos a teoria da classificação, a teoria do conceito e a teoria da comunicação. Com efeito, percebe-se que se trata de um campo que se dedica ao estudo de objetos complexos e, por conseguinte, tem como referências as práticas interdisciplinares. Nesse sentido, a presente unidade tem como objetivo discutir os principais elementos que subsidiam os processos de análise e representação da informação, no domínio interdisciplinar da Linguística, Terminologia, Lógica e Diplomática.

A interdisciplinaridade faz parte de uma rede conceitual de integração disciplinar, composta também pela pluridisciplinaridade, multidisciplinaridade e transdisciplinaridade. Cada uma dessas expressões designa um nível de inter-relação existente entre duas ou mais disciplinas (áreas de conhecimento), na abordagem de algum tema/problema de estudo. Aqui interessa saber que a interdisciplinaridade corresponde à colaboração entre duas ou mais disciplinas, que se caracteriza pelo compartilhamento mútuo de teorias, métodos e técnicas (SOUZA, 2015).

A análise da informação, conforme visto anteriormente, é uma atividade complexa que envida esforços práticos e, sobretudo, intelectuais. O aprimoramento dessa área de estudo vem se ampliando e exigindo cada vez mais a atuação colaborativa entre profissionais de diferentes áreas do conhecimento. Ela busca aportes teóricos em várias áreas, dentre as quais se podem destacar Ciência Cognitiva, Comunicação, Diplomática, Inteligência Artificial, Linguística, Lógica, Matemática e Terminologia.

A aproximação da análise da informação com diferentes áreas se deve, em grande medida, às diferentes formas de análise do texto que vêm sendo realizadas no interior de diferentes disciplinas. Nesse conjunto, *Cunha* (1987) destaca a análise de conteúdo, a análise literária, a análise semiótica e a análise linguística. Todas elas apresentam pontos de proximidade com a análise da informação, na medida em que têm finalidades idênticas, a saber, diferentes formas de expressar o conteúdo dos documentos.



Você sabia?

A transdisciplinaridade corresponde à aproximação de diferentes disciplinas e áreas do conhecimento visando ao compartilhamento de metodologias unificadoras e à articulação de métodos dessas áreas. Trata-se da produção de uma unidade do saber, na medida em que não existem fronteiras disciplinares. Na história da humanidade, não há um exemplo consolidado de transdisciplinaridade. A maior aproximação a esse empreendimento se deu na famosa *Escola de Sagres*, que reuniu astrônomos, cartógrafos, carpinteiros, engenheiros e matemáticos para a construção da caravela usada nas grandes navegações, movimento que resultou na descoberta das Américas e do Brasil. A *Escola de Sagres*, contudo, nunca existiu efetivamente, pois não havia prédio, estatuto, professores e nem mesmo alunos. A escola circulava entre a corte do rei de Portugal *D. Manuel (O venturoso)* e o cais do porto.

Fonte: DOMINGUES, I. (Org.). **Conhecimento e transdisciplinaridade II:** aspectos metodológicos. Belo Horizonte: Editora da UFMG, 2005.

Além disso, a interdisciplinaridade no domínio da análise da informação resulta, nas palavras de *Campos* (2001), da grande escassez de estudos teórico-metodológicos nas áreas de representação e recuperação da informação no Brasil. Essa situação evidencia a necessidade de estudos na área de Biblioteconomia e Documentação, mas também o aprofundamento de estudos já realizados em outras áreas do conhecimento, que permitam, a um só tempo, avançar na compreensão e no desenvolvimento de práticas de análise da informação e contribuir com outras áreas.

3.1 CONTRIBUIÇÕES DA LINGUÍSTICA E DA TERMINOLOGIA

A análise da informação foi realizada por muito tempo nos sistemas de recuperação da informação de modo intuitivo e empírico-pragmático, isto é, os processos de interpretação, seleção e síntese dos conteúdos informativos dos diversos documentos que compunham os acervos das bibliotecas dependeram da capacidade intuitiva e da experiência dos bibliotecários analistas. Porém, com a exigência dos modernos sistemas de informação, principalmente daqueles que ingressaram no processo de tratamento automático da informação, esses profissionais se viram obrigados a recorrer a diferentes procedimentos de análise de conteúdo já desenvolvidos em diferentes áreas do conhecimento.

A **Linguística** é um campo que se dedica ao estudo da língua natural e pode contribuir com a Biblioteconomia, a Documentação e, mais recentemente, a Ciência da Informação no domínio da análise e da síntese dos conteúdos informacionais e na elaboração de linguagens documentárias. Nesse sentido, *Smit* (1987) defende que a inter-relação entre análise da informação e Linguística se dá tanto no fornecimento de insumos básicos por parte desta àquela, quanto na validação de teorias da Linguística por meio de procedimentos realizados no domínio da análise documentária. Essa autora reconhece que a colaboração ainda não se faz de forma satis-

fatória porque, por um lado, as pesquisas linguísticas centram no espaço da sintaxe, consideradas insuficientes para a análise da informação, e, por outro, essa prática tem sido muito fundamentada na figura do analista com todas as suas idiosincrasias.

Na Linguística, uma das disciplinas que mais interessa é justamente a semântica, na medida em que trabalha com o processo de significação. Assim, embora *Cunha* (1987) reconheça que esta área contribua com a análise documentária no fornecimento de procedimentos analíticos e na elaboração de instrumentos auxiliares e de produtos, compreende que não existem efetivas construções interdisciplinares entre a Documentação e a Linguística. O fato é que a primeira utiliza diversos conhecimentos originários da segunda como subsídio teórico para sua prática analítica, mas parece não oferecer efetivas contribuições à área, que resultaria no trabalho colaborativo, próprio de empreendimentos interdisciplinares.

Não se pode perder de vista, contudo, que os aportes teóricos da Linguística são fundamentais para compreensão das estruturas das diferentes linguagens, que se constituem a partir de diferentes relações paradigmáticas e sintagmáticas. Com efeito, “[...] o reconhecimento das relações paradigmáticas e sintagmáticas leva não só à formulação de campos semânticos dos textos analisados como à estruturação e elaboração de linguagens tradutoras, permitindo a passagem de uma LN-LD” (CUNHA, 1987, p. 45).

Basta destacar que a diferenciação realizada por *Saussure* entre linguagem (capacidade humana de expressão e comunicação), língua (conjunto de códigos e regras) e fala (forma individual de comunicação) fornece à análise documentária embasamentos para a compreensão das diferenças entre linguagem natural e linguagem documentária. Porém, como campos de aplicação específicos da Linguística nos processos de análise e síntese da informação, podem ser destacados, entre outros, o **estruturalismo** de *Ferdinand de Saussure* (1857-1913), o **transformacionalismo** de *Avram Noam Chomsky* (1928-) e o **distribucionalismo** de *Zellig Sabbettai Harris* (1909-1992).

Conforme *Cunha* (1987), o desdobramento das propostas estruturalistas focaram no estudo de unidades mínimas da língua, isto é, em enunciados fora dos seus respectivos contextos de enunciação, que resultaram em descrições e taxonomias de conceitos operatórios com base em estruturas elementares.

Essa prática é bastante utilizada na elaboração de linguagens documentárias mais estruturadas, como as listas de cabeçalhos de assunto, que buscam, conforme abordado anteriormente, a definição de unidades de conhecimento e o estabelecimento de relações entre essas unidades, objetivando os controles semânticos e sintáticos.

O gerativismo transformacional tem como fundamentos as propriedades matemáticas das linguagens formais, na medida em que defendia a tese de que a gramática era necessariamente finita, mas tinha o papel fundamental de gerar um número infinito de frases. Nas palavras de *Cunha* (1987, p. 47), “dá-se, portanto, uma inversão em relação à proposta da gramática estrutural. Em lugar de uma análise, propõe-se uma síntese, manipulam-se sistemas de regras e não um sistema de elementos”.



Hans Peter Luhn

Nasceu no dia 1º de julho de 1896, na cidade de Barmen, na Alemanha, e faleceu, aos 68 anos, no dia 19 de agosto de 1964, nos Estados Unidos da América. Após ter concluído o ensino secundário, mudou-se para Suíça para aprender o ofício de impressão e trabalhar no negócio da família. Porém, após a Segunda Guerra Mundial, Luhn mudou-se para os EUA. Formado em Ciência da Computação, trabalhou na IBM, onde desenvolveu o denominado *algoritmo* Luhn e o índice **Key Words in Context** (KWIC), ingressando no campo da Ciência da Informação. Essa é uma das suas principais contribuições para a área de tratamento e recuperação da informação. Além destes relevantes trabalhos para a área de automação em sistemas de recuperação da informação, *Luhn* foi premiado com mais de 80 patentes. Foi presidente do **American Documentation Institute (ADI) / Association for Information Science and Technology (ASIS&T)**, em 1964, e recebeu prêmio de mérito desta instituição, no ano de 1974.

Fonte: LUHN, H. P. (1896-1964). **ASIS&T**: Association for Information Science and Technology, 2016. Disponível em: <<https://www.asist.org/pioneers/hans-peter-luhn/>>. Acesso em: 16 ago. 2016.

Nesse conjunto, um exemplo clássico da aplicação da Linguística na representação da informação foi o uso da inter-relação entre a Documentação e a Linguística distribucionalista de *Harris* realizada, desde os anos 1950, por *Hans Peter Luhn (1896-1964)* no desenvolvimento dos índices KWIC e KWOC. Vê-se que os estudos da Linguística tornaram-se fundamentais para a análise e síntese automáticas da informação.

O estudo do significado contribui para a pesquisa em processamento automático de texto. Com a evolução dessa automatização de textos, feita principalmente com a linguagem natural, é que foram iniciadas as pesquisas interdisciplinares, pois a utilização dessa linguagem traz sempre muita ambiguidade e inconsistência. A contribuição da linguística tem sido mais observada e desenvolvida, portanto, principalmente nos estudos da indexação automática. (DIAS; NAVES, 2007, p. 82).

Mais recentemente, estudos linguísticos e bibliométricos, acerca da função de índice temático da nominalização deverbal na escrita científica, desenvolvidos por *Guedes (2010)*, *Guedes, Mollica e Leal (2011)*, *Guedes e Santos (2013)*, entre outros, vêm intensificando a importância de abordagens teóricas e metodológicas das Ciências do Texto, especificamente da Morfologia, da Teoria de Gêneros Discursivos (BAZERMAN, 2005; HYLAND, 2009; SWALES, 1986 etc.), do Modelo de Frequência de Uso e de Tipo (BYBEE, 2007) e da Teoria Lexical (BASÍLIO, 1980), no contexto da Sociolinguística, para a identificação do conteúdo informativo de artigos científicos na Ciência da Informação. Nesse sentido, *Hjørland (2002)* aponta esses tipos de estudos linguísticos como abordagens relevantes para a análise de domínio na Ciência da Informação.

O estudo da *Terminologia* também se mostra indispensável ao domínio interdisciplinar da análise da informação, uma vez que o processo comunicacional que envolve os sistemas de informação e os seus respectivos usuários tem por base os diversos elementos que constituem as diferentes linguagens.

As pesquisas que abordam a terminologia permitem, além de compreender como se constituem e se relacionam os conceitos no domínio amplo da linguagem, entender como estes se inter-relacionam na produção e na comunicação científica. Essa questão toca de perto o campo da

representação da informação porque está diretamente relacionada ao estabelecimento de um domínio que aproxima áreas específicas e, de forma mais precisa, à elaboração de linguagens documentárias nelas usadas.

No espaço informacional verifica-se a necessidade de criação de instrumentos que possibilitem a comunicação, não mais entre os pares, mas entre os usuários de um sistema de informação e o próprio sistema, que seria o espaço do tesouro e da tabela de classificação. Porém, estes instrumentos, para serem criados, necessitam de uma estrutura terminológica que será buscada em um sistema terminológico (CAMPOS, 2001, p. 18).

A grande contribuição da Teoria Geral da Terminologia (TGT) para a análise da informação concerne aos fundamentos teórico-metodológicos acerca da constituição de sistemas de conceitos, que têm como objetivo último estabelecer um conjunto de conceitos representativos de uma área de conhecimento e as relações entre estes. Ao trabalhar com a noção de sistema, a definição de um conceito é realizada necessariamente a partir da relação que cada um deles estabelece com os demais. Há, a rigor, uma construção conceitual pautada no princípio da reciprocidade. Assim, a Terminologia tem natureza prescritiva, que é fundamental na elaboração e no uso das linguagens documentárias.

Ao procurar diferenciar Lexicologia e Terminologia, Campos (2001) esclarece que a primeira se instala no domínio da língua natural e, por conseguinte, trabalha com o universo caracterizado pela dispersão semântica e sintática. A segunda, por sua vez, atua no domínio da linguagem artificial, que busca construir, no espaço de um grupo de especialistas, a relação unívoca entre o conceito e a denominação.

Esta é uma questão que mostra claramente a relação da Terminologia com a evolução das linguagens documentárias, que representou um deslocamento da centralidade existente no “assunto”, como unidade de conhecimento, para o “termo”, como unidade de comunicação. Com efeito, “na TGT, o termo é a unidade de comunicação que representa o conceito e pode ser constituído de uma ou mais palavras, uma letra, um símbolo gráfico, uma abreviação, uma notação” (CAMPOS, 2001, p. 73).

Conforme apresentado anteriormente, as linguagens pré-coordenadas, consideradas tradicionais, usam assuntos simples, compostos e complexos para representar os conteúdos informativos dos documentos. Os assuntos são tomados, portanto, como unidades de conhecimento e se caracterizam essencialmente pela prescrição, pela enumeração e pela unidimensionalidade próprias dos sistemas pré-coordenados.

As linguagens pós-coordenadas passam a usar “termos”, que, determinados a partir de um sistema conceitual, designam um conceito. Este, portanto, corresponde ao significado do termo. Campos (2001, p. 71) esclarece que, “para a TGT, conceito é uma unidade de pensamento, constituído de características que refletem as propriedades significativas atribuídas a um objeto, ou a uma classe de objetos.”.

A integração de três teorias – Teoria da Classificação, Teoria Geral da Terminologia e Teoria do Conceito – traz grandes contribuições a este domínio interdisciplinar da análise da informação, na medida em que possibilita a leitura, a interpretação, a seleção e a síntese dos principais



conceitos que compõem os conteúdos dos documentos, em um contexto amplo e complexo de definição de categorias, relações e domínios.

3.2 CONTRIBUIÇÕES DA LÓGICA

A **lógica** é um ramo da filosofia que estuda as formas de pensamento e as diferentes operações intelectuais que envolvem o raciocínio e a argumentação, tais como dedução, hipótese, indução e inferência, entre outras. As operações lógicas são realizadas como método dedicado a estabelecer as diferentes relações entre duas ou mais proposições.

Conforme *Dias e Naves* (2007), dois tipos de operações lógicas são usadas na análise textual: o **raciocínio dedutivo** (silogismo aristotélico) e o **raciocínio indutivo** (inferência probalística). A **dedução** é um **método não ampliativo** a partir da relação entre duas premissas verdadeiras, obtém-se necessariamente uma conclusão verdadeira – “todo A é B” e “C é A”, então “C é B”. A dificuldade neste tipo de raciocínio analítico se encontra no fato de existir algumas deduções cujas premissas maiores são condicionais (“Se A é B”) e não partem necessariamente de premissas gerais, como acontece nos *modus tollens* e *ponens*. Além disso, é preciso considerar a possibilidade de uma das premissas não ser verdadeira.

De base estatística, a **indução**, ao contrário do raciocínio anterior, é um **método ampliativo**, que parte de casos particulares e alcança uma verdade geral – de uma amostra de “A”, em que 80% estão grafados na fonte “N”, conclui-se que todos os “A” estão grafados na fonte “N”. A dificuldade no uso deste método analítico é que a validade dos resultados depende da representatividade da amostra utilizada na construção do raciocínio e se caracteriza como uma probabilidade, na medida em que o método estatístico é a base de sua sustentação.

Observa-se, pois, que os procedimentos lógicos são fundamentos importantes para a realização de processos interpretativos e seletivos, que compreendem a base da análise da informação. Essa área de conhecimento passou a ser usada como aporte teórico-metodológico nos processos de análise e síntese da informação, desde o momento em que se percebeu que os problemas da análise documentária não se restringiam aos aspectos semânticos e sintáticos, mas incluíam também o texto em sua estrutura e dinâmica.

Há, nas palavras de *Smit* (1987), uma ampliação da abordagem que ultrapassa o universo frásico inicial e alcança a lógica interna do texto, que é igualmente relevante para a análise. Observe que, nessa mudança de perspectiva analítica, foram ultrapassadas algumas barreiras:

- a) a estrutura frásica não é necessariamente uma unidade relevante de informação, uma vez que uma informação relevante pode estar presente em palavra, frase, período, parágrafo ou no próprio texto;
- b) nos casos em que a unidade de análise passa a ser o próprio texto, o seu gerenciamento está condicionado ao domínio da sua estru-

tura interna, requerendo, portanto, do analista a valorização das relações lógico-semânticas, presentes em todo o texto e

- c) ao centrar as operações analíticas de um documento no próprio texto, ele deixa de ser um espaço neutro, na medida em que se passam a considerar as condições de produção e de uso do texto. Com efeito, o texto deixa de ser um elemento isolado e fechado, e passa a ser um espaço de circulação de informações, que, por sua natureza, são carregadas de acréscimos realizados pelos diversos atores envolvidos no processo de produção, ou seja, os autores, o analista da informação e os usuários.

Em meio a esse conjunto de mudanças, duas questões sobressaem. Primeiro, o texto que compõe os diversos documentos não é neutro ou fechado em si mesmo e, portanto, não está condicionado a uma leitura única e absoluta. Segundo, assim como autor e usuário, o analista da informação tem um papel ativo nos processos de interpretação e seleção de conceitos.

A lógica, em última instância, se apresenta como importante instrumento para o analista da informação porque, além de auxiliar no desenvolvimento do raciocínio lógico, com o estabelecimento de premissas e inferências, fundamenta a interpretação e a construção de uma rede de relações entre os diferentes conceitos que compõem o conteúdo informacional dos documentos.

Na condição de analista da informação, você deve considerar que “[...] a análise documentária passa a ser uma modalidade de leitura do texto que leva em conta, além do próprio texto, as condições de produção e consumo” (SMIT, 1987, p. 9). E o que efetivamente isso significa?

Do ponto de vista prático, pelo menos duas situações em que o uso da lógica é fundamental na análise e síntese da informação. Primeiro, as informações dos elementos objetos da leitura documentária são incompletas ou imprecisas. Segundo, o que é mais grave, o próprio conteúdo informacional se mostra bastante incompleto e/ou impreciso, exigindo maior esforço por parte do analista na interpretação e definição do domínio conceitual do documento. Em outros termos, com o uso da lógica, o analista é obrigado a correlacionar, nas palavras de *Cunha* (1989), o **catálogo de conhecimentos** e o **catálogo de raciocínios**, na interpretação e na seleção de conteúdos informativos.

Quadro 6 – Categorias de análise lógica

Categoria analítica	Descrição
Agente	Sujeito que pratica a ação
Paciente	Sujeito que sofre a ação
Objeto	Motivo da ação
Instrumento	Os meios concretos que possibilitam a ação
Modo	Forma como se pratica a ação
Espaço	Lugar onde a ação foi praticada
Tempo	Momento em que a ação foi praticada
Produto	Resultado da ação
Finalidade	Intenção do sujeito na prática da ação

Fonte: Adaptado de CUNHA (1989).



Cabe, então, considerar a definição de categorias analíticas e, complementarmente, a descrição e o estabelecimento de relações, como exemplificado no Quadro 6, que auxiliam o analista no desenvolvimento de raciocínios lógicos e de juízos acerca dos conteúdos informacionais.

3.3 CONTRIBUIÇÕES DA DIPLOMÁTICA

A origem da **Diplomática** está relacionada à determinação da autenticidade dos documentos e remonta ao período medieval, mas as primeiras tentativas de sistematização dos estudos do **diploma** ocorreram já no século XVII, com a publicação da primeira edição da obra *De re diplomatica libri sex*, em 1681, de *Jean Mabillon*, conhecido como *Dom Mabillon*. Esta obra se apresenta como um tratado de Diplomática em que foram delimitados seu objeto de estudo e seu campo de atuação, centrados no documento administrativo-jurídico, que, em síntese, se caracteriza pela sua forma solene de produção.

O tipo de documento a ser considerado, portanto, objeto de estudo da Diplomática clássica, dependerá de seus objetivos, diante da estrutura construída por meio de fórmulas solenes de validação no sistema social-jurídico-acadêmico à época. Portanto, para que um documento possa ser objeto de estudo da Diplomática clássica, é preciso analisar seu objetivo de prova ou testemunho no contexto funcional e/ou da administração pública ou privada (NASCIMENTO, 2009, p. 112).

Há na delimitação do objeto de estudo da Diplomática clássica a preocupação com a precisão da natureza do documento diplomático, que se caracteriza pelas condições de sua escritura, pela natureza jurídico-administrativa do fato comunicado e pela forma adotada na compilação do documento.

Considerando esse entendimento, apenas uma parcela de documentos é de interesse para a análise diplomática. Porém, contemporaneamente, alguns estudiosos vêm fazendo um exercício que visa ao melhor esclarecimento e à ampliação da noção de documento diplomático, embora permanecendo como centralidades de análises a estrutura formal e a autenticidade dos documentos.

A Diplomática é, pois, outro campo científico que mantém proximidades conceituais e procedimentais com a análise da informação, na medida em que ambas se dedicam, em maior ou menor medida, à identificação e análise da estrutura e dos conteúdos informativos dos documentos. Neste contexto, *Nascimento* (2009) aponta as bases que as aproximam, de um lado, as **formas diplomáticas** e a **superestrutura do texto** (tipologia textual) e, de outro, o **conteúdo diplomático** e as **macroestruturas do texto** (conteúdo informacional).

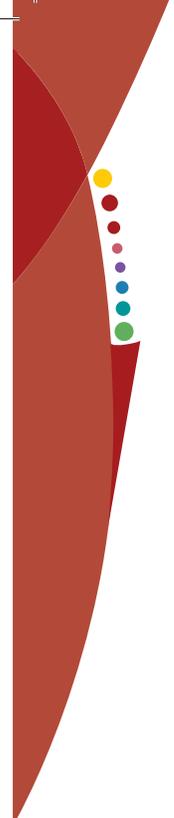
Quadro 7 – Estrutura de análise diplomática de um livro

Categorias	Elementos	Enunciados conceituais
Elementos Externos	Suporte	Tipo de papel
		Impressão
		Cor
	Escritura	Sumário
		Parágrafo
		Índice
		Gráficos
		Listas
	Linguagem	Dissertativo-argumentativa
		Manual de estudo
		Texto prático
	Selos e anotações	Código de barra
		<i>International Standard Book Number (ISBN)</i>
		Logomarca
		Tiragem
Elementos internos	Protocolo inicial	Titulação
		Destinatário
		Agradecimento
		Dedicatória
	Textos e subseções	Preâmbulo
		Anúncio da obra
		Exposição dos motivos
		Assunto tratado
		Direitos autorais
	Escatocolo	Editora
		Local de publicação
		Data de publicação
		Citações de outras fontes
		Referências às fontes

Fonte: Adaptado de NASCIMENTO (2009).

Tomando como referência a diversidade que caracteriza o universo dos documentos, tanto da superestrutura quanto das macroestruturas, é compreensível que em determinadas circunstâncias os procedimentos adotados nas análises diplomáticas se tornem, mais ou menos, adequadamente aplicáveis ao domínio da análise da informação. Esta é uma constatação que pode ser feita a partir do exercício analítico realizado por Nascimento (2009), que teve como base material os livros *Senso crítico: do dia-a-dia às Ciências Humanas* (CARRAHER, D. W., 1993) e *A fome*





de ler (BARKER, R. E. ; ESCARPIT, R., 1975). A autora analisou o primeiro documento, considerando os elementos externos e internos, conforme ilustrado no Quadro 7.

A partir do estudo de *Nascimento* (2009), algumas considerações podem ser realizadas. Primeiro, cada documento exige uma análise circunstanciada. Segundo, a estrutura analítica adotada pode ser adaptada de forma relativamente fácil à análise e síntese do conteúdo informacional de qualquer documento. Terceiro, os enunciados conceituais dos elementos internos são os mais importantes para a análise da informação, em função de sua materialidade (conteúdo) e formalidade (arranjo). Por fim, quarto, os enunciados conceituais externos têm papel imprescindível na compreensão do domínio conceitual e da garantia literária, elementos importantes na aproximação necessária entre documento e usuário.

Há na literatura uma diversidade de aplicações do método diplomático na análise de informação, em diferentes documentos, tais como atas de reunião, documentos eletrônicos, manuais técnicos, plantas baixas, receitas culinárias e rótulos de xampu, entre outras. As experiências já realizadas demonstram que a tipologia documental tem relação direta com a modalidade de texto, que evidencia uma forma de expressão, mas mantém também similaridades com alguns elementos da macroestrutura.

Em que pesem as particularidades daquela área de conhecimento no que se refere à preocupação com as funções administrativas e jurídico-comprobatórias dos documentos, os conceitos e os procedimentos nela desenvolvidos são importantes fundamentos e instrumentos para análise da informação. Embora tenha uma tradição como ciência auxiliar nos domínios da História e, mais tardiamente, da Arquivística, a Diplomática se apresenta atualmente como importante área de estudos e práticas colaborativas para a Biblioteconomia e a Documentação, principalmente no domínio da análise das macroestruturas dos conteúdos informativos que compõem os documentos. Esse conjunto de procedimentos tem como fundamento o **método diplomático**, que tem como objetivo a interpretação e a representação do conteúdo dos documentos em diferentes contextos.

Há uma aproximação entre a Diplomática e a Documentação, constituindo um domínio pluridisciplinar de interesse comum da Arquivologia e da Biblioteconomia. Desse modo, considera-se, a partir de *Guimarães e Rabello* (2007, p. 144), que “[...] a Diplomática pode estabelecer uma profícua relação interdisciplinar na aplicação de seu método para a análise e identificação do conteúdo temático do documento em ambiências distintas (biblioteca e arquivos)”.

O método diplomático desempenha importante função na definição de estratégias metacognitivas de leitura e interpretação dos conteúdos informativos dos documentos, na medida em que define estruturas físicas e previsíveis de documentos como espaço de articulação entre forma e conteúdo. Mais que isso, a Diplomática amplia a noção de pragmática nos processos de análise e síntese dos conteúdos informacionais, ao estabelecer um nexos entre os diversos tipos de documentos e os seus respectivos usos. O fato é que, do ponto de vista prático, a Diplomática centra suas análises na tipologia documental e nas instituições que expediram os documentos analisados.

Os procedimentos da Diplomática se destacam por suas características descritivas e pelas diferentes possibilidades de aplicação, considerando a

preocupação em dar cobertura às particularidades estrutural e funcional dos diversos tipos de documentos. Nesse sentido, a análise diplomática centra basicamente em três domínios, a saber, na tipologia documental, na procedência dos documentos e nos elementos extrínsecos e intrínsecos.

O estilo do texto, que compreende as construções semânticas e sintáticas dos diversos documentos, constitui-se em importante elemento de análise e síntese dos conteúdos informativos, na medida em que define o domínio conceitual e reflete o contexto em que foi produzido.

Em que pese o fato de não existir, na literatura técnico-científica, construções teórico-metodológicas suficientes que demonstrem relações interdisciplinares entre análise documentária e análise diplomática, a ampliação e o aprofundamento dos estudos informacionais evidenciam as diversas possibilidades de compartilhamento e colaboração entre estas duas subáreas, notadamente na adoção de procedimentos sistemáticos de interpretação e seleção dos elementos que exprimem os conteúdos informativos dos documentos, que apontam para a integração de saberes neste domínio de estudo.



RESUMO

A análise da informação objetiva a construção de informações documentárias que possibilitam o tratamento, a organização e a recuperação dos conteúdos informativos dos documentos. Os estudos sobre análise e síntese têm como base uma rede teórico-conceitual que requer a adoção de práticas colaborativas a partir de relações interdisciplinares entre a Biblioteconomia, a Documentação e a Ciência da Informação e diferentes áreas do conhecimento, tais como Linguística, Terminologia, Lógica e Diplomática, entre outras. A rigor, observa-se que as relações interdisciplinares se atêm à recorrência a essas áreas do conhecimento, sem evidentes interações recíprocas. A **Linguística** é dedicada ao estudo da língua natural e contribui com os estudos sobre análise e síntese dos conteúdos e com a elaboração de linguagens documentárias. Nessa área, especificamente os estudos semânticos são fundamentais para a compreensão das diferentes estruturas de linguagens por meio de relações paradigmáticas e sintagmáticas. A **Terminologia**, complementarmente, é indispensável à compreensão da comunicação presente nos sistemas de recuperação da informação, na medida em que ela possibilita o entendimento de como os conceitos se constituem e se interrelacionam no domínio da linguagem, na definição de campo semântico e na produção e comunicação científica, bem como na construção de sistemas de conceitos, fundamentais à elaboração e ao uso de linguagens documentárias. A **Lógica**, por sua vez, estuda as formas de pensamento e as operações intelectuais que envolvem raciocínio e argumentação. No domínio da análise da informação, as operações lógicas são bastante relevantes porque fundamentam os processos interpretativos e seletivos, que possibilitam o alcance da dinâmica interna dos conteúdos informativos dos documentos. Trata-se de importantes contribuições para a análise e síntese da informação porque auxilia no raciocínio lógico, com o estabelecimento de premissas e

inferências, na interpretação e na construção de redes de relações entre diferentes conceitos que compõem o campo semântico do sistema de recuperação da informação. E, por fim, a **Diplomática**, que historicamente se dedica à precisão da natureza do documento diplomático, compreende as condições de escritura, a natureza jurídico-administrativa do fato documentado e a forma de compilação do documento. Observamos sua relação com a análise da informação no domínio da identificação e da análise da estrutura e dos conteúdos dos documentos. O método diplomático é, pois, imprescindível na definição de estratégias metacognitivas de leitura e interpretação porque define as estruturas do documento como espaço de articulação entre forma e conteúdo, considerando sua tipologia e seu uso.



3.4 Atividade

1. Explique fundamentadamente a necessidade de conhecimentos e práticas interdisciplinares no processo de análise da informação.

2. Os estudos interdisciplinares entre Biblioteconomia, Documentação e Linguística podem resultar em importantes contribuições para a área de análise da informação. Nesse contexto, é correto afirmar que:
 - a) As pesquisas na Linguística centram na semântica, campo essencial à compreensão e definição da ordem de citação, nos estudos documentários.
 - b) A análise da informação é um domínio dedicado ao fornecimento de insumos básicos para os estudos da Linguística.
 - c) A Linguística fornece contribuições para a análise da informação e para a construção de linguagens documentárias, especialmente no domínio da semântica.
 - d) A Linguística corresponde ao campo de validação das teorias e práticas desenvolvidas na Documentação e na Biblioteconomia.
 - e) A efetiva colaboração entre Biblioteconomia, Documentação e Linguística, no campo da análise da informação, é amplamente reconhecida.

3. A Terminologia é uma área de conhecimento que fornece subsídios importantes aos estudos e às práticas de análise da informação. Analise as assertivas abaixo:
- I. Os estudos terminológicos no domínio da linguagem documentária resultaram no deslocamento da noção de “assunto” para a noção de “termo”.
 - II. O princípio da reciprocidade estabelece a assistemática de na definição dos conceitos informacionais.
 - III. A Terminologia é fundamental na compreensão dos conceitos e na definição de domínios e relações conceituais.
 - IV. Os estudos terminológicos no domínio da linguagem documentária resultaram no deslocamento da “unidade de comunicação” para a “unidade de conhecimento”.

Estão corretas:

- a) I e II, apenas.
 - b) I e III, apenas.
 - c) II e III, apenas.
 - d) III e IV, apenas.
 - e) I, II, III e IV.
4. A Lógica é um ramo da Filosofia que tem como objeto de estudo as formas de pensamento e as operações intelectuais. Nesse sentido, pode oferecer importantes contribuições aos estudos de análise da informação. Leia as afirmações abaixo e marque a correta:
- a) O raciocínio dedutivo é um tipo de operação lógica usado na análise textual, que tem como fundamento a estatística e a matemática.
 - b) A indução é um método ampliativo, uma vez que obtém uma conclusão verdadeira a partir de premissas gerais também verdadeiras.
 - c) A dedução é um método restritivo que se operacionaliza a partir de inferências probalísticas.
 - d) A indução é um método ampliativo e probalístico, pois parte de uma amostra e obtém conclusões universais.
 - e) A adoção de operações lógicas, no domínio da análise da informação, impulsionou a exclusão das condições de produto e do uso do texto nos processos analíticos.
5. O método diplomático teve historicamente como objetivo a análise e a determinação da autenticidade dos documentos jurídico-administrativos. A ampliação da Diplomática, contudo, possibilitou a melhor aplicação deste à análise documental. Leia as assertivas abaixo:
- I. A aplicação do método diplomático na análise da informação está relacionada à ampliação da noção de documento na Diplomática e à manutenção da análise na estrutura formal e na autenticidade dos documentos.
 - II. A linguagem é definida pelo método diplomático como um importante elemento analítico da macroestrutura do texto.
 - III. O texto e as suas respectivas subseções compreendem a superestrutura analítica de um documento.



IV. O assunto tratado e as subseções constituem importantes elementos para análises pautadas nas macroestruturas dos documentos.

Estão corretas:

- a) I e II, apenas.
- b) I e III, apenas.
- c) I e IV, apenas.
- d) I, II, III e IV.
- e) II e III, apenas.

REFERÊNCIAS

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **NBR 6023**: informação e documentação – referências. Rio de Janeiro: ABNT, 2018.

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **NBR 6028**: informação e documentação – resumos. Rio de Janeiro: ABNT, 2003.

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **NBR 6034**: informação e documentação – índice. 2. ed. Rio de Janeiro: ABNT, 2004. 4 p.

BASÍLIO, M. M. P. **Estruturas lexicais do português**: uma abordagem gerativa. Petrópolis: Editora Vozes, 1980. 128 p.

BAZERMAN, C. **Gêneros textuais, tipificação e interação**. São Paulo: Cortez, 2005.

BERTALANFFY, L. V. **Teoria geral dos sistemas**: fundamentos, desenvolvimentos e aplicações. Petrópolis, RJ: Vozes, 2008. 360 p.

BYBEE, J. L. **Frequency of use and the organization of language**. New York: Oxford University Press, 2007.

CAMPOS, M. L. A. **Linguagem documentária**: teorias que fundamentam sua elaboração. Niterói, RJ: Ed UFF, 2001. 133 p.

CAMPOS, M. L. A.; GOMES, H. E.; MOTTA, D. F. Elaboração de tesauro documentário: tutorial. **Biblioteconomia, informação & tecnologia da informação** jul. 2004. Disponível em: <<http://www.conexaorio.com/bit/tesauro/>>. Acesso em: 11 nov. 2015.

CARNEIRO, M. V. Diretrizes para uma política de indexação. **Revista da Escola de Biblioteconomia da UFMG**, Belo Horizonte, v. 14, n.

2, p. 221-241, set. 1985. Disponível em: <<http://portaldeperiodicos.eci.ufmg.br/reb/>>. Acesso em: 11 nov. 2015.

CESARINO, M. A. Sistema de Recuperação da informação. **Revista da Escola de Biblioteconomia da UFMG**, Belo Horizonte, v. 14, n. 2, p. 157-168, set. 1985. Disponível em: <<http://portaldeperiodicos.eci.ufmg.br/reb/>>. Acesso em: 26 dez. 2015.

CESARINO, M.; PINTO, M. C. M. F. Cabeçalho de assunto como linguagem de indexação. **Revista da Escola de Biblioteconomia da UFMG**, Belo Horizonte, v. 7, n. 2, p. 268-288, set. 1978. Disponível em: <<http://portaldeperiodicos.eci.ufmg.br/reb/>>. Acesso em: 11 nov. 2015.

CINTRA, A. M. M. Estratégias de leitura em documentação. In: SMIT, J. W. (Coord.). **Análise documentária: a análise da síntese**. Brasília: IBICT, 1987. p. 27-35.

CINTRA, A. M. M. et al. **Para entender as linguagens documentárias**. 2. ed. São Paulo: Polis, 2002. 92 p.

CUNHA, I. M. R. F. Análise documentária. In: SMIT, J. W. (Coord.). **Análise documentária: a análise da síntese**. Brasília: IBICT, 1987. p. 38-60.

CUNHA, I. M. R. F. "O falcão maltês": a lógica em análise documentária. **Revista de Biblioteconomia de Brasília**, v. 17, n. 1, p. 51-61, jan./jun. 1989. Disponível em: <http://www.brapci.ufpr.br/brapci/_repositorio/2011/07/pdf_5636de33b8_0017680.pdf>. Acesso em: 11 nov. 2015.

DEWEY, M. **Dewey Decimal Classification**. Dublin, OH: OCLC, 2003. 4 v.

DIAS, E. J. W.; NAVES, M. M. L. **Análise de assunto: teoria e prática**. Brasília: Thesaurus, 2007. 116 p.

DODEBEI, V. L. D. **Tesouro: linguagens de representação da memória documentária**. Rio de Janeiro: Interciência, 2002. 119 p.

DOMINGUES, I. (Org.). **Conhecimento e transdisciplinaridade II: aspectos metodológicos**. Belo Horizonte: Editora da UFMG, 2005.

FUNDAÇÃO BIBLIOTECA NACIONAL. **Catálogo**. Rio de Janeiro, 2015. Disponível em: <http://acervo.bn.br/sophia_web/index.html>. Acesso em: 11 nov. 2015.

FUJITA, M. S. L. A leitura documentária na perspectiva de suas variáveis: leitor-texto-contexto. **Data grama zero – revista de Ciência da Informação**, v. 5, n. 4, ago. 2004. Disponível em: <http://www.brapci.inf.br/_repositorio/2010/01/pdf_af744c7a63_0007646.pdf>. Acesso em: 03 nov. 2018.

GIL LEIVA, I.; FUJITA, M. S. L. (Ed.). **Política de indexação**. São Paulo: Cultura Acadêmica; Marília: Oficina Universitária, 2012. 260 p. Disponível em: <http://www.culturaacademica.com.br/_img/arquivos/Livro%20Politica-de-indexacao_ebook.pdf>. Acesso em: 11 nov. 2015.



GROSSE, E. U. **Text und Kommunikation**: eine Linguistische Einführung in die Funktionen der Texte. Stuttgart: Kohlhammer, 1976.

GUEDES, V. L. S. **Nominalizações deverbais em artigos científicos**: uma contribuição para a análise e a indexação temática da informação. Rio de Janeiro, 2010. 109 f. Tese (Doutorado em Linguística) – Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2010. Disponível em: <<http://www.lettras.ufrj.br/poslinguistica/wp-content/uploads/2013/03/vania-lisboa.pdf>>. Acesso em: 11 nov. 2015.

GUEDES, V. L. S.; MOLLICA, M. C. M.; LEAL, M. B. B. Produtividade léxico-morfológica em artigos científicos em língua portuguesa numa perspectiva interdisciplinar. In: CONGRESO Internacional de la Asociación de Linguística y Filología de la América (ALFAL), 16, 2011, Madrid. **Actas...** Madrid: Universidad de Alcalá de Henares, 2011.

GUEDES, V. L. S.; SANTOS, M. J. V. C. Recorrência de nominalizações deverbais em resumos de cartas científicas em língua portuguesa e a indexação temática. **Linguística**, v. 29, n. 1, p. 37-57, jun. 2013. Disponível em: <<http://www.scielo.edu.uy/pdf/ling/v29n1/v29n1a03.pdf>>. Acesso em: 11 nov. 2015

GUIMARÃES, J. A. C.; RABELLO, R. A contribuição metodológica da diplomática para a análise documental de conteúdos em arquivos e bibliotecas. In: RICHTER, E. I. S.; ARAUJO, J. C. G. (Org.). **Paleografia e diplomática no curso de arquivologia UFSM**. Santa Maria, RS: FACOS – UFSM, 2007. p. 137-157.

HANS PETER LUHN. **ASIS&T**: Association for Information Science and Technology, 2016. Disponível em: <<https://www.asist.org/pioneers/hans-peter-luhn/>>. Acesso em: 16 ago. 2016.

HJØRLAND, B. Domain Analysis in Information Science: Eleven Approaches – Traditional as Well as Innovative. **Journal of Documentation**, v. 58, n. 4, p. 422-462, 2002.

HYLAND, K. **Academic Discourse**: English in a Global Context. New York: Continuum International Publishing Group, 2009. 215 p.

KATO, M. Uma visão interativa da legibilidade. **Ilha do Desterro**: a Journal of English Language, Literatures in English and Cultural Studies, Florianópolis, v. 13, p. 57-66, 1985. Disponível em: <<https://periodicos.ufsc.br/index.php/desterro/article/view/10533/10052>>. Acesso em: 11 nov. 2015.

LANCASTER, F. W. **Indexação e resumo**: teoria e prática. 2. ed. Brasília: Briquet de Lemos Livros, 2004. 452 p.

LANGRIDGE, D. **Classificação**: abordagem para estudantes de Biblioteconomia. Rio de Janeiro: Interciência, 2006. 120 p.

MAC DEVELOPER LIBRARY. **Searchkit Programming Guide**. 2005. Disponível em: <<https://developer.apple.com/library/mac/documentation/UserExperience/Conceptual/SearchKitConcepts/>>

searchKit_intro/searchKit_intro.html#/apple_ref/doc/uid/TP40002842-TPXREF101>. Acesso em: 11 nov. 2015.

MODESTO, F. **Metadados**: introdução básica. São Paulo: Versão, 2005. (Apostila de aula da disciplina de Representação Descritiva II). Disponível em: <<http://www2.eca.usp.br/prof/fmodesto/disc/RDII/texto/metadado2005f.pdf>>. Acesso em: 05 dez. 2015.

NASCIMENTO, L. M. B. **Análise documental e análise diplomática**: perspectivas de interlocução de procedimentos. Marília, SP. 198 f. (Doutorado em Ciência da Informação) – Universidade Estadual Paulista, Marília, 2009. Disponível em: <https://www.marilia.unesp.br/Home/PosGraduacao/CienciadaInformacao/Dissertacoes/nascimento_lmb_do_mar.pdf>. Acesso em: 11 nov. 2015.

OTLET, P. **Traité de Documentation**. Le livre sur le livre. Théorie et pratique. Bruxelles: Éditeurs-Imprimeurs D. van Keerberghen & Fils, 1934.

PIEDEDE, M. A. R. **Introdução à teoria da classificação**. 2. ed. Rio de Janeiro: Interciência, 1983. 221 p.

PINTO, M. C. M. F. Análise e representação de assuntos em sistemas de recuperação de informação; linguagens de indexação. **Revista da Escola de Biblioteconomia da UFMG**, Belo Horizonte, v. 14, n. 2, p. 169-186, set. 1985. Disponível em: <<http://portaldeperiodicos.eci.ufmg.br/reb/>>. Acesso em: 11 nov. 2015.

RANGANATHAN, S. R. **Prolegomena to Library Classification**. Bombay: Asia Publishing House, 1967. 640 p.

SMIT, J. W. Introdução. In: SMIT, J. W. (Coord.). **Análise documentária**: a análise da síntese. Brasília: IBICT, 1987. p. 5-10.

SOUZA, E. D. **A Ciência da Informação**: fundamentos epistêmico-discursivos do campo científico e do objeto de estudo. Maceió, Alagoas: Edufal, 2015. 222 p.

SOUZA, E. D. **A epistemologia interdisciplinar**: uma introdução à produção colaborativa de conhecimento científico. Maceió: Edufal, 2015. 106 p.

SOUZA, E. D. Configurações do campo da Ciência da Informação: pluralismo epistemológico e descentração interdisciplinar. **Tendências da Pesquisa Brasileira em Ciência da Informação**, v. 5, n. 1, 2012. Disponível em: <<http://inseer.ibict.br/ancib/index.php/tpbci/article/view/63/104>>. Acesso em: 15 set. 2015.

SWALES, J. M. Citation analysis and discourse analysis. **Applied Linguistics**, v. 7, n. 1, p. 39-56, 1986.

VERGUEIRO, W. **Seleção de materiais de informação**: princípios e técnicas. 3. ed. Brasília: Briquet de Lemos, 2010. 120 p.

WERLICH, E. **Typologie der Texte**. Heilderberg: Quelle und Meyer, 1975.



