

# 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.

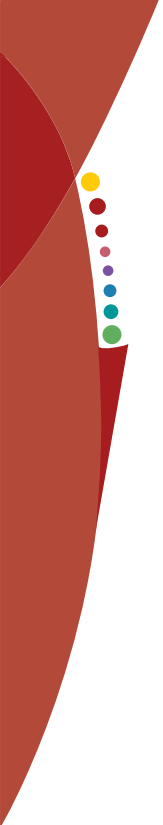
## 2.1 LINGUAGEM NATURAL E LINGUAGEM ARTIFICIAL

---

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:

<b>Information Retrieval</b> x Recuperação da Informação	<b>Information Science</b> x Ciência da Informação	<b>recall</b> 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 (**filo**)

Chelicerata (**subfilo**)

Arachnida (**classe**)

Scorpiones (**ordem**)

Euscorpidae (**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)

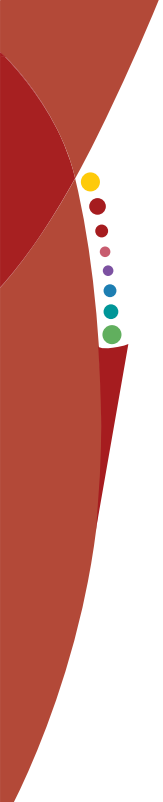
Exemplo de assunto complexo:

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**

<b>Instrumento</b>	<b>Notação</b>
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**

<b>Instrumento</b>	<b>Notação</b>
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, 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.

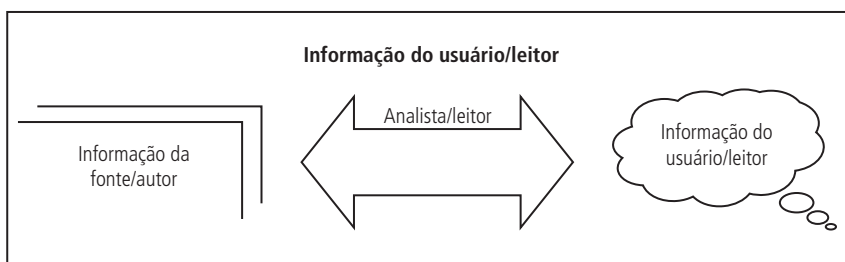
Conforme Kato (1985), há dois tipos de estratégias de leitura: **cognitivas** e **metacognitivas**. As primeiras são realizadas de forma inconsciente e automática pelo leitor. As segundas se dão a partir de comportamentos conscientes e, portanto, controlados do leitor. A autora considera, contudo, que 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**

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

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**

<b>CDD</b>	
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**

<b>Categoria analítica</b>	<b>Temática</b>
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	<b>020.905</b>

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 <b>ver</b> BIBLIOTECONOMIA – HISTÓRIA	Teorias da classificação <b>ver</b> CLASSIFICAÇÃO – TEORIA	Tipos de resumos <b>ver</b> 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	BIBLIOTECA PÚBLICA TG Biblioteca
TR Arquivologia Biblioteca Centro de informação Museu	TR Biblioteca Escolar Biblioteca Especial Biblioteca Especializada Biblioteca Universitária
ARQUIVOLOGIA TR Arquivo Biblioteconomia Ciência da Informação Documentação Museologia	BIBLIOTECA UNIVERSITÁRIA TG Biblioteca
BIBLIOTECA TE Biblioteca Escolar Biblioteca Especial Biblioteca Especializada Biblioteca Pública Biblioteca Universitária	TR Biblioteca Escolar Biblioteca Especial Biblioteca Especializada Biblioteca Pública
TR Arquivo Centro de Informação Museu	BIBLIOTECONOMIA TR Arquivologia Ciência da Informação Documentação Museologia
BIBLIOTECA ESCOLAR TG Biblioteca	CENTRO DE INFORMAÇÃO TR Arquivo Biblioteca Museu
TR Biblioteca Especial Biblioteca Especializada Biblioteca Pública Biblioteca Universitária	CIÊNCIA DA INFORMAÇÃO TR Arquivologia Biblioteconomia Documentação Museologia





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:

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

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**

<b>A</b>	
Arquivo, 5, 8, 15, 20, 25, 36	
<b>Ver também</b> 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>B</b>	
Biblioteca, 40, 43-45, 52, 60, 61, 63- 64, 75	
<b>Ver também</b> 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	
	[...]
<b>Z</b>	
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.

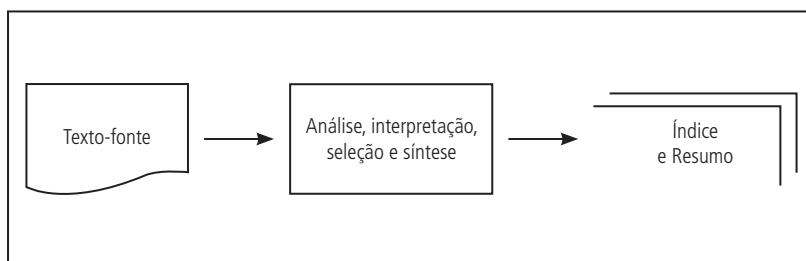
Existem também índices que se constituem em verdadeiros documentos referenciais. Essas são publicações analíticas elaboradas por serviços especializados de indexação e resumo, que assumem o formato de resumos (*abstracts*). Como exemplos, destacam-se o *Chemical Abstracts* (CA) e o *Library and Information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

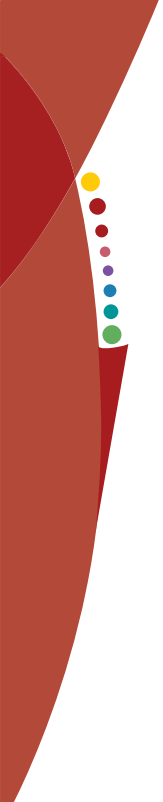
**Figura 16 – Índice e resumo: materialidade e formalidade**



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.

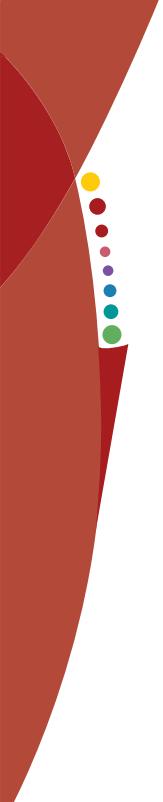
**Figura 23 – Resumo altamente formatado**

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 consultantes, 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.

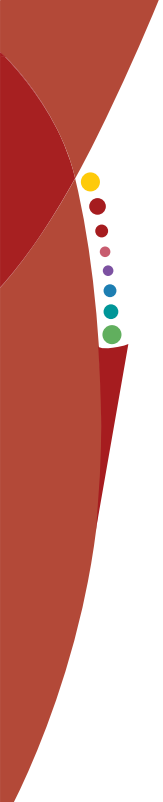
Essa primeira aproximação da base material das análises deve considerar a política de indexação e elaboração de resumos que define as medidas de extensão da representação da informação quanto à abrangência semântica (geral ou específica) e à cobertura do conteúdo do documento (seletiva ou exaustiva). 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 documental, 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, extrínseca 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 (cabeçalhos 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 cabeçalhos 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-





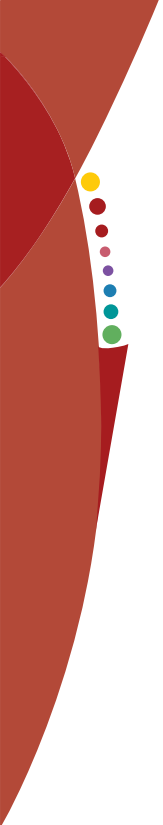


- I. A indexação por extração, conhecida também como indexação por conceitos, é realizada usando a linguagem natural.
- II. A indexação por atribuição, conhecida também por indexação por palavras, é realizada usando a linguagem artificial.
- III. A indexação por atribuição, conhecida também como indexação 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 afinitivas são tipos de relações ampliativas, que ofertam termos alternativos na recuperação da informação.

Estão corretas:

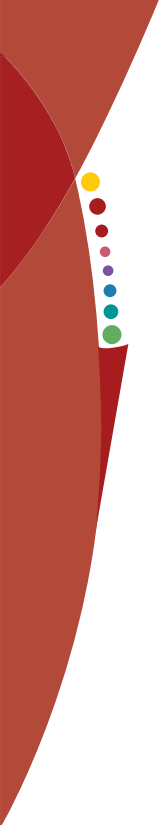
- a) I e II, apenas.
  - b) I e III, apenas.
  - c) II e IV, apenas.
  - d) III e IV, apenas.
  - e) I, II, III e IV.
6. A coordenação é uma estratégia usada para especificar a representação da informação. Analise as alternativas e aponte a correta.
- a) Na pré-coordenação, o usuário do sistema especifica 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.
-