



Tesouro



Profa. Lillian Alvares,
Faculdade de Ciência da Informação, Universidade de Brasília

Dicionário Analógico de Peter Mark Roget,
"Thesaurus of English words and phrases", 1852





Histórico

Anos 1940 e 1950

- ▶ **Crescimento** da produção científica
- ▶ Necessidade de acelerar o processo de **disseminação** do conhecimento produzido.
- ▶ Em decorrência, nos anos **50, a recuperação de informação ganha força entre a comunidade científica.**
 - ▶ **Os tesouros vieram como resposta a essa necessidade de recuperação de informação.**

Antes disso...

- ▶ A palavra tesouro tem origem do latim **thesaurus**, que significa **tesouro**.
- ▶ Foi empregada como título no Dicionário Analógico de **Peter Mark Roget**, "Thesaurus of English words and phrases", publicado em Londres pela primeira vez em 1852.
 - ▶ Trabalhou nesse projeto durante **50 anos**.

-
- ▶ Em seu dicionário as palavras foram **agrupadas** em ordem **distinta da alfabética**.
 - ▶ Priorizaram-se as **ideias** que exprimiam e esta foi a ordem escolhida.
 - ▶ A busca por palavras dava-se sempre por aquilo que elas podiam expressar, com seu **significado**.

-
- ▶ A formação de um tesouro dá-se por palavras cuidadosamente escolhidas, palavras que **possuam significado para uma determinada área.**
 - ▶ Na área de documentação, associou-se à forma de **organização do vocabulário para os processos de indexação e recuperação**

Brunetto Latini

- ▶ Alguns autores afirmam que o termo tesouro já havia sido utilizado anteriormente...
- ▶ ... por **Brunetto Latini** (1220-1294) para designar uma **enciclopédia** sistematizada que ele chamou de ...

□ Os Livros do Tesouro

Etienne

- ▶ Em 1532, **Thierry e Robert Etienne** publicaram o livro (Tesouro do idioma latino' (**Thesaurus Latinde Linguae**) um dicionário da língua latina em arranjo alfabético.
- ▶ Quarenta anos depois, Henri Etienne II: **Thesaurus language Graecae**, iniciado por seu pai R. Etienne.

-
- ▶ Em 1736, Shorter Oxford Dictionary registrou o uso da expressão inglesa "thesaury or storehouse of knowledge“...
 - ▶ ...definindo o termo como ...
 - ▶ ... **"tesouro ou armazém de conhecimento, similar a um dicionário ou a uma enciclopédia"**

No Brasil

▶ Destaque para Prof. Francisco Ferreira dos Santos Azevedo

- ▶ Entendia que as palavras tinham **vínculos não apenas segundo uma sinonímia**, isto é, num sentido estreito, particular, mas sim de forma mais **ampla**, ou seja, de acordo com **ideias básicas**, com grupos de ideias afins.

- Passou a viver, por mais de **quinze anos**, no mundo das palavras, **procurando conhecê-las melhor**, conhecer suas **origens** e seus diferentes **graus de parentesco**.
- Dessa convivência resultou uma obra monumental da semântica,
 - **Dicionário Analógico da Língua Portuguesa em 1950**

Contexto de surgimento

Voltando...

- ▶ O tesouro surgiu da necessidade de **manipulação de grande quantidade de documentos especializados...**
 - ▶ ... onde é preciso trabalhar com vocabulário mais específico e uma estrutura mais articulada e integrada.

Inovações

▶ O tesouro avançou na...

▶ Na **estrutura** e

▶ nas **referências cruzadas**

□ dando lugar às **relações hierárquicas (verticais) e associativas (horizontais)**.

Inovações

- ▶ Essas relações entre os termos ...
 - ▶ conferem a um tesouro sua características e...
 - ▶ ... uma multiplicidade de usos...
 - ...desde o auxílio na **função de indexação** até
 - o auxílio na efetiva **recuperação de informação**.

No momento da indexação

- ▶ No momento da indexação, por meio dos processos consecutivos de:
 - ▶ **análise** do documento
 - ▶ **identificação** de seu conteúdo
 - ▶ **tradução** para os termos do tesauro de acordo com a política de indexação.

No momento da recuperação

- ▶ Na escolha dos termos

NISO Z39.19

▶ *National Information Standards Organization*

▶ <http://www.niso.org/home/>

▶ Ambiente para definição de compromissos estratégicos colaborativos, por meio de normalização, para o desenvolvimento e utilização de normas e para os serviços de informação.



NISO Z39.19

- ▶ Tesouro é um **vocabulário controlado** organizado em uma **ordem preestabelecida** e **estruturado de modo que os relacionamentos entre os termos de:**
 - ▶ **equivalência**
 - ▶ **de homografia**
 - ▶ **de hierarquia**
 - ▶ **de associação**
 - ▶ sejam indicados claramente e identificados por indicadores de relacionamento padronizados.

NISO Z39.19

- ▶ As finalidades primordiais de um tesouro são
 - ▶ **facilitar a recuperação dos documentos** e
 - ▶ alcançar a consistência na **indexação** dos documentos

NISO Z39.19

- ▶ Instrumento de controle terminológico
 - ▶ ... utilizado para **traduzir a linguagem natural** dos:
 - ▶ **documentos,**
 - ▶ **Indexadores**
 - ▶ **usuários**
 - **para uma linguagem sistêmica mais contida**

Definição 1

- ▶ Lista **estruturada de termos associados** empregada por indexadores para...
- ▶ ... **descrever um documento** com a desejada especificidade e para permitir aos usuários a recuperação da informação que procuram.

Definição 2

- ▶ *No contexto da organização e recuperação da informação, tesouros são instrumentos de controle terminológico, **utilizados em sistemas de informação** para traduzir a linguagem dos documentos, dos indexadores e dos pesquisadores numa linguagem controlada, usada na indexação e na recuperação de informações.*

Tipologia

- ▶ Os tesouros são classificados quanto à:
 - ▶ **Tipologia:** de acordo com a(s) língua(s) que englobam, em monolíngues e multilíngues.
 - ▶ **Conceitos:** macrotesouros (representam conceitos mais amplos) e microtesouros (representam conceitos específicos)
 - ▶ **Abrangência temática:** multidisciplinares ou de disciplina específica.

Vantagens

- ▶ Capacidade de **descrever as informações de forma completa.**
- ▶ **Auxiliam o indexador** a localizar o conceito num campo semântico **e o usuário** a identificar assuntos relacionados que podem ser de interesse na busca.
- ▶ O emprego das **novas tecnologias** da informação possibilita **diferentes formas de apresentação, mais amigáveis e compreensíveis para orientar os usuários na navegação,** ampliando ou restringindo sua busca.

Principais finalidades

- ▶ **Controlar os termos**
- ▶ **Uniformizar os termos** em uma instituição ou numa rede cooperativa.
- ▶ **Limitar o número de termos necessários à explicitação** dos conceitos expostos pelos autores de uma área.
- ▶ Auxiliar a tarefa de **recuperação da informação**, fornecendo termos adequados para a estratégia de busca.



Normas para elaboração

- ▶ ISO 2788 (1986)
- ▶ ISO 5964 (1985), no caso de tesouro multilíngue
- ▶ ANSI / NISO Z39.19-2003, que define diretrizes de construção e formatação de tesouros.
- ▶ Demais diretrizes e manuais de elaboração de tesouros, **disponíveis também em língua portuguesa.**

Tesouro x Dicionário

- ▶ Theasaurus usualmente:
 - ▶ Não contém informações sobre
 - Etimologia
 - Pronúncia
 - Uso
 - Gramática

Campos do Tesauro

- ▶ **Descritor:** termo escolhido para representar um conceito e que será utilizado na indexação e na recuperação de determinado assunto.
 - *Quando houver outros termos que representem o mesmo conceito, antes do termo descritor, constará a sigla USE.*
- ▶ **Não-descritor:** Termo que, embora descreva o mesmo conceito que o descritor, não é autorizado na indexação, para evitar a proliferação de sinônimos. Antes de cada não-descritor, constará a sigla UP.

Campos do Tesouro

- ▶ **Nota explicativa (NE):** Fornece uma definição do termo ou uma orientação sobre como utilizá-lo em uma indexação.
 - ▶ É facultativo
- ▶ **Categoria (CAT):** Grande grupo ao qual pertence o descritor

Campos do Tesauro

- ▶ **Termo genérico (TG):** Indica que há relação hierárquica entre termos com relação gênero-espécie e que este descritor representa o termo com o conceito mais abrangente.
- ▶ **Termo específico (TE):** Indica os termos subordinados ao termo genérico na cadeia hierárquica.
- ▶ **Termo relacionado (TR):** Indica relação entre termos que não formam uma hierarquia (gênero-espécie). Servem para orientar o *indexador quanto às possibilidades de encadeamento de descritores* e para *sugerir ao usuário formas de limitar ou expandir uma pesquisa*.

Exemplo

- ▶ Frutas
 - ▶ TG Produtos agrícolas
 - ▶ TE Frutas Cítricas

- ▶ Frutas Cítricas
 - ▶ TG Frutas
 - ▶ TE Laranja
 - ▶ TE Limão

- ▶ Laranja
 - ▶ TG Frutas Cítricas

- ▶ Produtos Agrícolas
 - ▶ TE Frutas

Características

- ▶ Os tesouros são construídos **para uma área específica do conhecimento.**
- ▶ Não existe um tesouro **geral; alguns cobrem vários assuntos.**

Características

- ▶ Deve **incorporar as mudanças que a linguagem** utilizada sofre no caminho de seu desenvolvimento sem abrir mão dos conceitos, mas em atitude aberta a seu próprio desenvolvimento.

Objetivos

- ▶ A estrutura de um tesouro é que traz ao usuário - **consulente ou indexador** - a possibilidade de encontrar o termo mais adequado, mesmo desconhecendo o nome específico para a idéia ou conceito.
- ▶ A partir do que informa ao usuário, **outros termos, oportunos ou não, são apresentados para sua escolha.**





Relacionamentos

-
- ▶ A estrutura de um tesouro diz respeito **aos relacionamentos, às ligações e vinculações existentes entre os conceitos** representados por termos.
 - ▶ Os relacionamentos podem ser de diversos tipos, como:

Tipos

- ▶ Relacionamentos Lógicos (ou hierárquicos)
- ▶ Relacionamentos Ontológicos
- ▶ Relacionamentos de Efeito
- ▶ Relacionamentos de Equivalência

Relacionamentos Lógicos

- ▶ Dividem-se em:
 - ▶ Genérico-específico
 - ▶ Analítico (ou associativo)
 - ▶ De Oposição
 - ▶ Enumerativa



Relacionamientos Lógicos

Que permite formar as classes dos conceptos

- ▶ Animal
 - ▶ Mamíferos
 - ▶ Elefantes
 - Elefantes indianos
 - Elefantes indianos de Kerala

Genérico-específico



Relacionamentos Lógicos

- ▶ Que gera as relações a partir das **relações associativas entre os termos**
 - ▶ Animal
 - ▶ Defesa dos meio ambiente
 - Proteção de mamíferos ameaçados
 - Extinção do sagui-preto-de-mão-amarela

Analítico
(ou associativo)

Relacionamentos Lógicos

- ▶ Indicando os termos opostos
 - ▶ Grande
 - ▶ Pequeno
 - ▶ Caro
 - ▶ Barato
 - ▶ Limpo
 - ▶ Sujo

De Oposição

Relacionamentos Lógicos

- ▶ Relaciona vários itens coordenados
 - ▶ Mar
 - Mar Báltico
 - Mar Negro
 - Mar Mediterrâneo

Enumerativa



Relacionamentos Ontológicos (ou Partitivo)

▶ Reunindo as relações partitivas ou o todo e suas partes.

▶ Automóvel

□ Sistema de Alimentação

□ Tanque de combustível

Relacionamentos de Efeito

- ▶ Constituído pelos relacionamentos de causalidade ou **causa e efeito** e pelos os relacionamentos de **descendência** que apresentam relações genealógicas entre os termo.
- ▶ Debilidade Renal
 - ▶ Hepatite
 - Supressão de alimentos selecionados

Relacionamentos de Equivalência

- ▶ Sinonímia ou Quase-Sinonímia
 - ▶ São aquelas que representam os sinônimos ou quase sinônimos de um termo (mesma intenção).
 - ▶ Linguagens de documentação
 - ▶ Linguagens de indexação

Relacionamentos de Equivalência

▶ Homônímias

- ▶ A homonímia é o emprego do mesmo termo com significados diferentes. A solução, quando ocorre, é a indicação do contexto.

- Tênis (calçado)

- Tênis (esporte)

Síntese dos Relacionamentos

- ▶ Relacionamentos Lógicos (ou hierárquicos)
 - ▶ Genérico-Específico
 - ▶ Analítico (ou associativo)
 - ▶ De Oposição
 - ▶ Enumerativo
- ▶ Relacionamentos Ontológicos (ou Partitivo)
- ▶ Relacionamentos de Efeito
- ▶ Relacionamentos de Equivalência
 - ▶ De Sinonímia ou Quase-Sinonímia
 - ▶ De Homonímia

Conclusão dos relacionamentos

- ▶ As **relações** existentes entre os termos de um tesouro constituem **ponto fundamental** de tal forma que as diretrizes de **alguns tesouros proíbem termos órfãos**, isto é, não relacionados a outros termos.
- ▶ Por outro lado, em **circunstâncias em que é o único acesso** à informação é necessário incluir termos órfãos ou conceitos importantes para que possam ser recuperados.

Planejamento de Tesouros

Planejamento

1. Delimitação da Área
2. Público Alvo
3. Classificação
4. Levantamento das Fontes
5. Forma de Apresentação
6. Período de Atualização
7. Divulgação
8. Seleção do Software
9. Manutenção



I. Delimitação da área

- ▶ O assunto deve ser bem delimitado, também porque um número muito grande de conceitos é de difícil sistematização e aumenta a possibilidade de sinonímias.
- ▶ Se o assunto for amplo, é necessário recortá-lo em assuntos mais específicos, estruturados como microtesauros.

2. Público alvo

- ▶ O tesouro não possa ser útil a todos igualmente.
 - ▶ Dependendo da abordagem, as relações entre os conceitos podem se alterar.
 - ▶ O universo de conceitos abrangidos vai variar no grau de especificidade.

3. Classificação

- ▶ Um exercício prévio de classificação é fundamental para estabelecer os limites do tema do tesouro.

4. Levantamento das Fontes

- ▶ O perfil do usuário é determinante para a seleção das fontes:
 - ▶ Os relatórios técnicos de projetos podem ser uma fonte interessante
 - ▶ Dicionários especializados
 - ▶ Índices de publicações periódicas
 - ▶ Revistas técnicas

5. Forma de Apresentação

- ▶ A tradicional forma de apresentação é alfabética.
- ▶ Se a consulta for feita apenas em computador, a forma alfabética não tem sentido....

Planejamento

- ▶ UP - Usado Para
- ▶ USE - Use
- ▶ NE - Nota explicativa
- ▶ CAT - Categoria
- ▶ TG - Termo Geral
- ▶ TE - Termo Específico
- ▶ TR - Termo Relacionado

Forma de
Apresentação
Tradicional

Planejamento

- ▶ Termos:
 - ▶ Termo imediatamente superior (TG)
 - ▶ Termo(s) imediatamente inferior(es) (TE)
 - ▶ Termo relativo ao todo (TGP)
 - ▶ Termo(s) relativos às partes/componentes/constituintes (TEP)
 - ▶ Termo(s) associado(s) (TA)

Forma de
Apresentação

6. Período de Atualização

- ▶ A atualização de um tesouro deve ser uma atividade permanente.
- ▶ A prática de registro permanente de ocorrência de novos termos e definições, com discussão permanente entre os responsáveis pelo tesouro, torna a atividade terminológica mais eficaz e menos penosa.

7. Divulgação

- ▶ Antes: quando se pensava apenas na edição impressa
- ▶ Hoje: ampla divulgação em todos os suportes

8. Seleção do software

- ▶ Características necessárias:
 - ▶ Impedir a entrada de um termo mais de uma vez
 - ▶ Possibilitar a criação de diversos tipos de relacionamento
 - ▶ Produzir relatórios hierárquicos
 - ▶ Produzir relatórios alfabéticos, com as relações
 - ▶ Possuir dispositivo para alterações

9. Manutenção do Software

- ▶ Periodicamente é necessário atualizar o vocabulário e, de forma permanente, acompanhar o desenvolvimento do software. Um grupo deve ser constituído especificamente para esse fim.



Garantia literária, garantia de usuário,
garantia de uso e garantia estrutural

Bases para a construção de um tesouro

- ▶ Um dos principais desafios no processo de construção de um tesouro refere-se à:
 - ▶ seleção de vocabulário, que vem logo após as atividades de planejamento do tesouro, ou seja,...
 - ... a definição do domínio
 - dos objetivos do instrumento
 - da seleção das fontes para consulta e
 - a definição de sua forma de apresentação.



Bases para a construção de um tesouro

- ▶ É importante limitar o tamanho de um vocabulário para se ter somente termos necessários e suficientes para alcançar seus objetivos.

Garantia Literária

- ▶ Nas linguagens de indexação, o vocabulário escolhido para a representação dos assuntos deve ser derivado da literatura para a qual pretende-se a representação.
- ▶ Ou seja, a literatura deve ser determinante.

Garantia do Usuário

- ▶ Os termos selecionados para um tesouro precisam estar de acordo com aqueles utilizados pelos usuários na tarefa de recuperação de informação.
 - ▶ Alguns teóricos da indexação por assuntos vêm a garantia do usuário competindo em importância com a garantia literária.

Garantia de Uso

- ▶ Muitas vezes, alguns termos não se apóiam nem na garantia literária nem na garantia de usuário, mas são admitidos em um vocabulário normalizado porque possibilitam uma função estrutural útil.
 - ▶ São termos que facilitam elos em uma hierarquia de termos ou colaboram para que seja possível dispor um conjunto mais específico de termos.

Garantia Estrutural

- ▶ Quando se constrói um vocabulário controlado, o mais provável é admitir termos que estão somente garantidos em vocabulários altamente estruturados, como os empregados por linguagens classificatórias.

Referências

- ▶ Moreira, M.P.; Moura, M.A. **Construindo tesouros a partir de tesouros existentes: a experiência do TCI - Tesouro em Ciência da Informação.** DataGramZero - Revista de Ciência da Informação, v.7, n.4, ago/06.

