

# Recuperação

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

Library  
Information



search ID: rman57

"I CAN'T FIND THE BOOKS ON  
INFORMATION RETRIEVAL."

# Contexto

- **Grande quantidade** de informações são produzidas e disponibilizadas diariamente
- Com a elevada quantidade de documentos, o processo de **recuperação de informação não é uma tarefa simples**
- A baixa qualidade associada a boa parte dos documentos disponíveis **prejudica a recuperação de informações relevantes.**

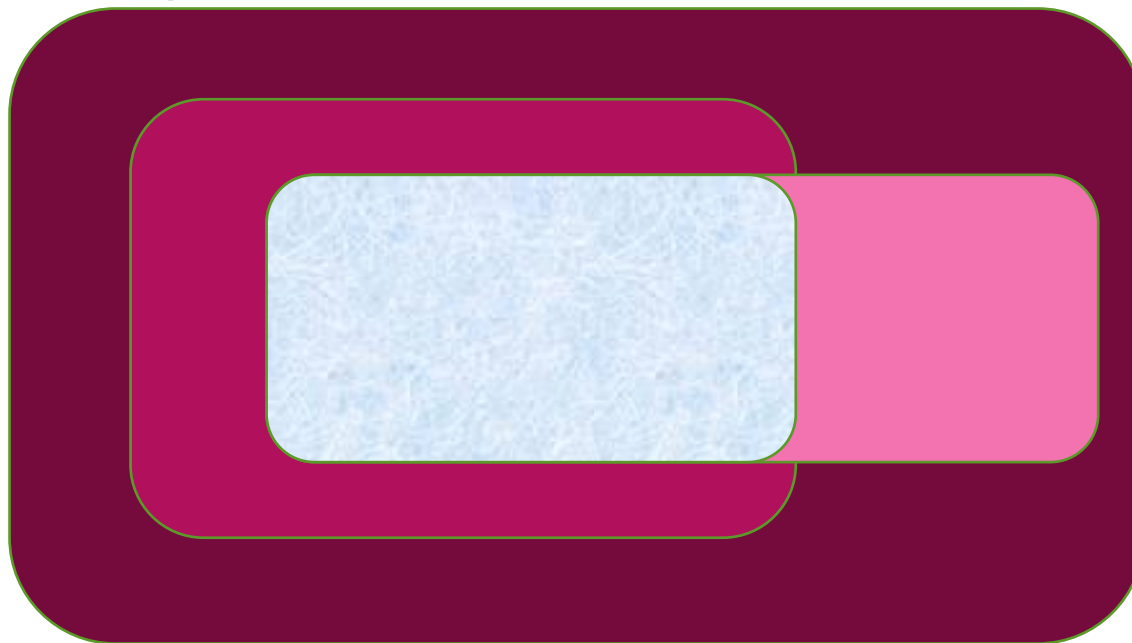
# Definição

- Processo de **localizar documentos** que tenham sido objeto de armazenamento, com a finalidade de permitir o **acesso dos usuários** aos itens de informação, objetos de uma solicitação

# Medidas de Eficácia


---

# Revocação e Precisão



	<b>Todos os documentos</b>
	Documentos relevantes
	Documentos retornados
	Relevantes retornados



  
= Revocação

  
= Precisão

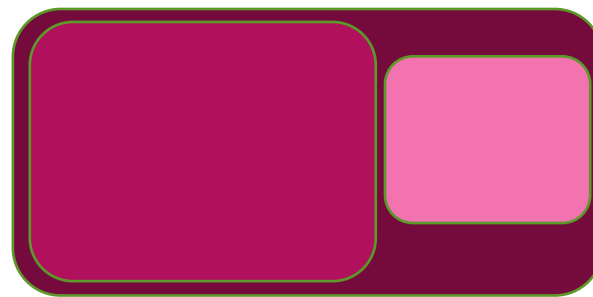
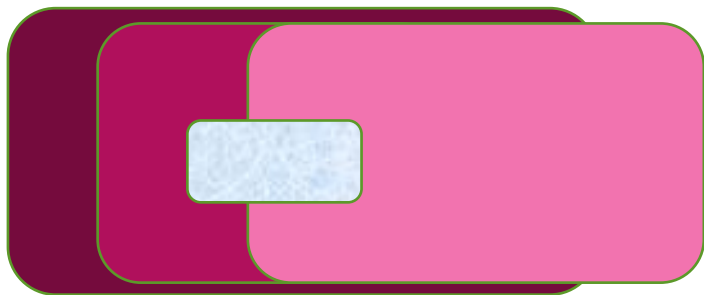
 + 

# Revocação (Recall)



 = Revocação  


# Precisão



	<b>Todos os documentos</b>
	Documentos relevantes
	Documentos retornados
	Relevantes retornados

$$\frac{\text{Relevantes retornados}}{\text{Documentos retornados}} = \text{Precisão}$$

A diagram showing a pink square followed by a plus sign and a light blue square, representing the denominator of the precision formula.



# Alguns Elementos da

---

# Recuperação de Informação

# Base de Dados

- **Base de dados** do Sistema de Recuperação de Informação

# Indexação

- Processo de **criação de índices** para os documentos que serão adicionados à base de dados

# Operações Textuais

- São operações realizadas sobre o texto de modo que **seja reduzido o texto a ser pesquisado**, a fim de que o processo de recuperação seja mais eficiente.
- Consistem em **retirar artigos, conjunções, preposições, acentos e caracteres especiais e preposições** que fazem elo entre as palavras.

# Interface

- Estrutura responsável para permitir a **interação** do usuário com os SRI

# Pergunta

- São **operações** realizadas sob expressão ou palavra-chave utilizada pelo usuário para que o SRI possa fazer a busca.

# Busca

- Estrutura que consiste em **buscar** na base de dados as informações

# Ranking

- Consiste em **classificar os documentos recuperados na Busca** por ordem decrescente ou crescente de relevância e apresentá-los ao usuário.



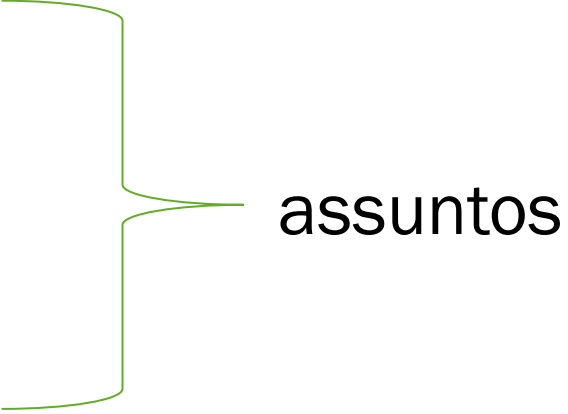
# Operadores Booleanos

---

## Operadores Lógicos

### Busca Lógica

# Operadores Booleanos

- São palavras ou símbolos utilizados para:
    - combinar
    - ampliar
    - excluir
- 
- assuntos

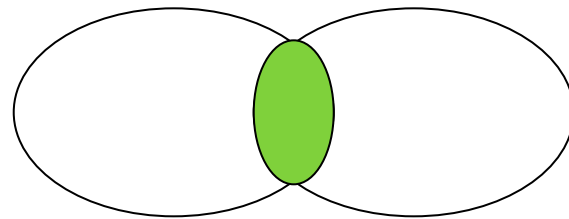
# Operadores Booleanos

OPERADOR	FUNÇÃO	RESULTADO
AND (E)	Combinar	<b>dort AND men</b> (Disfunções Osteomusculares Relacionadas ao Trabalho E em homens)
OR (OU)	Ampliar	<b>dort OR ler</b> (Disfunções Osteomusculares Relacionadas ao Trabalho OU Lesões por Esforço Repetitivo)
NOT (NÃO)	Excluir	<b>dort NOT men</b> (Disfunções Osteomusculares Relacionadas ao Trabalho NÃO em homens)

# Operadores Booleanos

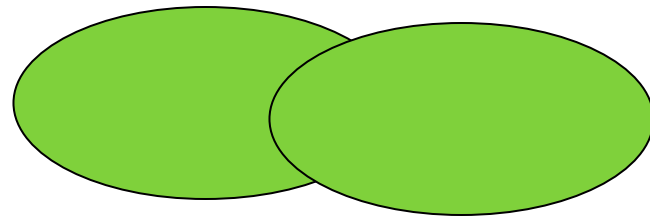
- Operadores lógicos

- and (intersecção)

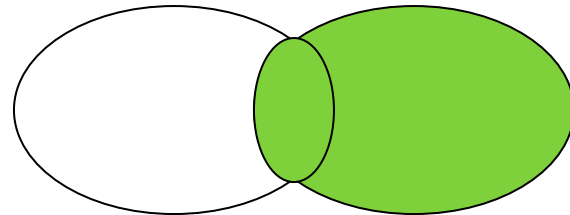


- 

- or (união)



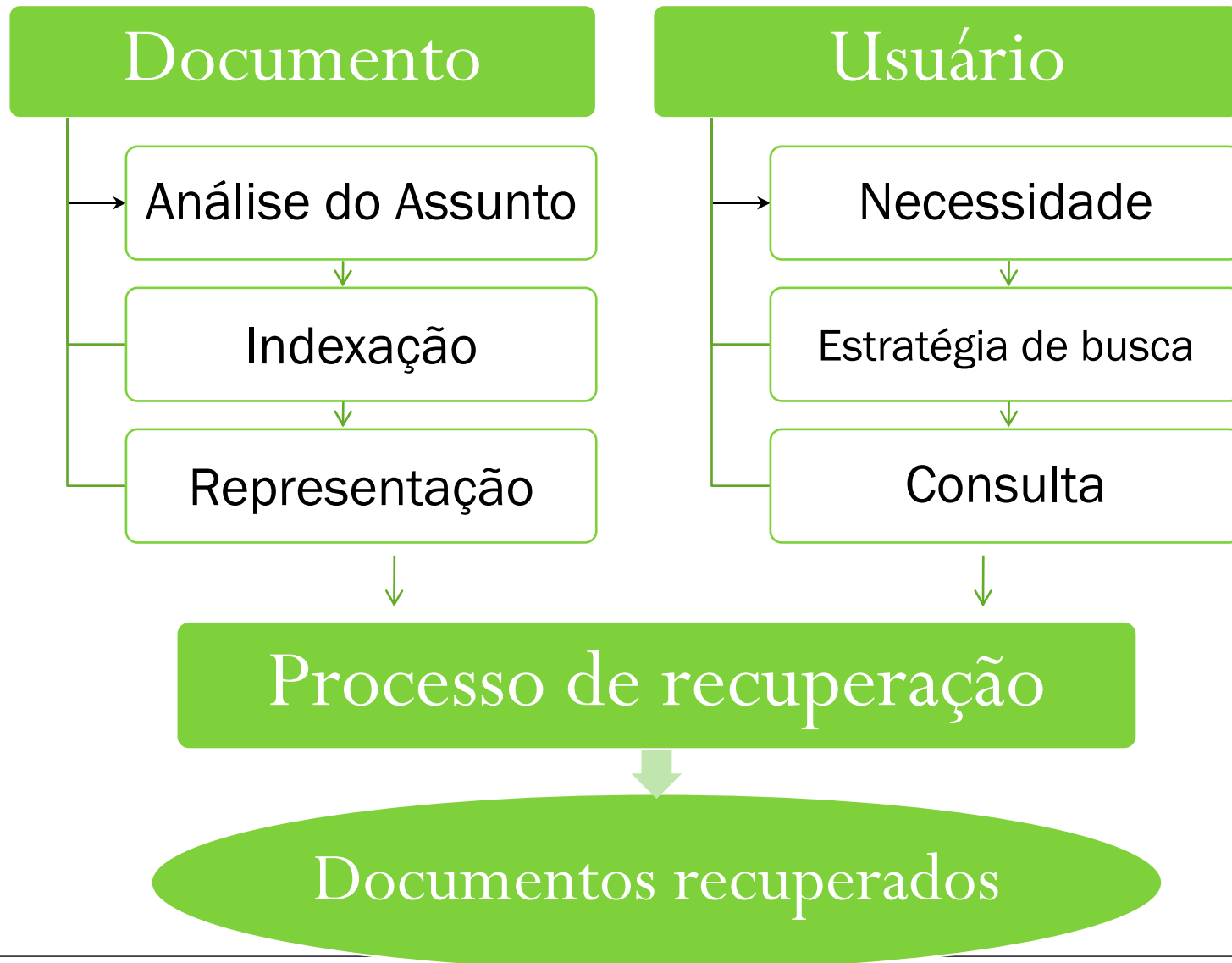
- not (exclusão)



# Truncamento de palavras

- ?
  - ?logia recupera:
    - biologia
    - geologia
    - psicologia
    - etc

# Sistema de Recuperação da Informação



# Usuário

---

# Tipologia

- Novato:
  - aquele que **nunca usou** determinado sistema e prefere **interfaces simples** e **intuitivas**



# Tipologia

- Ocasional:
  - utiliza o sistema de modo **infrequente e precisa reaprender a usá-lo cada novo contato**

# Tipologia

- Especialista:
  - possui **conhecimento** da coleção ou da base de dados e é **familiarizado** com a interface do sistema

# Tipologia

- Experiente:
  - usa o sistema frequentemente e **pode lidar com problemas** que venham a surgir

# Tipologia

- Com **necessidades especiais**:
  - pode ser portador de deficiências visuais ou auditivas, ou com necessidades físicas especiais ou distúrbios de compreensão

# Recuperação e Internet

---

# Recuperação e Internet

- Web:
  - Imenso repositório de documentos
    - **Motores de busca** (search engines)

# Recuperação e Internet

- Web semântica:
  - pretende embutir inteligência na confecção de páginas web, de modo a melhorar a forma com que podem **interagir** com essas páginas e também possibilitar um uso mais intuitivo por parte dos usuários

# Metadados:

---

## datos sobre datos



# O que são Metadados?

- Conjunto de dados estruturados que **identificam os dados de um determinado documento. São dados sobre o documento.**

# O que são Metadados?

- São informações **úteis para identificar, localizar, compreender e gerenciar os dados.**

# O que são Metadados?

- **Representam informações que caracterizam a informação documentada. .**

# Utilização

- Metadados se aplicam a uma **grande variedade de acervos:**
  - Bibliotecas
  - Sistemas de Informações Geográficas
  - Objetos
  - Multimídia, etc.

# Referências

- ARAÚJO JR., Rogério Henrique de. **Precisão no processo de busca e recuperação da informação**. Brasília : Thesaurus, 2007. 176 p.
- GUINCHAT, C.; MENOU, M. **Introdução geral às ciências e técnicas da informação e documentação**. 2. ed. Brasília: IBICT, 1994. 540p.
- LOPES, Ilza Leite. Estratégia de busca na recuperação da informação: revisão da literatura. **Ciência da Informação**, Brasília, v.31, n.2, p.60-71, maio/ago. 2002. Disponível em: <http://revista.ibict.br/index.php/ciinf/article/view/158/137>. Acesso em: 23 set. 2011.