

*Institut Supérieur d'Informatique et des Mathématiques de Monastir*



# **TECHNIQUES D'INDEXATION ET RECHERCHE MULTIMÉDIA**

**IMEN CHEBBI**

# Chapitre 3

**1. Formulation des  
requêtes**

**2. Modèles de RI**

# REQUÊTES PAR MOTS CLÉS

- L'utilisateur exprime ses besoins d'information avec une liste de (au moins un) mots-clés (ou termes) visant à trouver des documents contenant certains (au moins un) ou tous les termes de la requête.
- Les termes dans la liste sont supposés être connectés avec une version "douce" de l' ET logique.



- ✓ Par exemple, si l'on s'intéresse à la recherche d'informations sur 'Web mining, on peut émettre la requête « **Web mining** » sur un SRI.
- ✓ «**Web mining**» se retire en tant que «Web » et « mining». Le SRI trouve alors les documents susceptibles d'être pertinents pour les présenter à l'utilisateur.



- Notez qu'un document récupéré ne doit pas contenir tous les termes dans la requête.
- **Dans certains SRI**, l'ordre des mots dans la requête est également significative et affectera les résultats de recherche.



# REQUÊTES BOOLÉENNES

- L'utilisateur peut utiliser les opérateurs booléens, AND, OR et NOT pour construire des requêtes complexes.
- Ainsi, ces requêtes se composent de termes et d'opérateurs booléens.
- Par exemple, 'data OR Web' est une requête booléenne, qui demande des documents contenant le mot 'data' ou 'Web.'



# REQUÊTES BOOLÉENNES

- Une page est renvoyée pour une requête booléenne si la requête est logiquement vraie dans la page (c.-à-d. Correspondance exacte).
- Bien que l'on puisse écrire des requêtes booléennes complexes à l'aide des trois opérateurs, les utilisateurs écrivent rarement de telles requêtes.



# REQUÊTES BOOLÉENNES

Les moteurs de recherche acceptent généralement une version restreinte des requêtes booléennes.



# REQUÊTES PAR PHRASES

- Consiste en une séquence de mots qui constitue une phrase.
- Chaque document retourné doit contenir au moins une instance de la phrase.

# REQUÊTES PAR PHRASES

- ❖ Dans un moteur de recherche, une requête de phrase est normalement incluse avec des guillemets.
- ❖ Par exemple, on peut émettre la requête de phrase suivante, "Web mining techniques and applications" pour trouver des documents contenant la phrase exacte.

# REQUÊTES DE PROXIMITÉ

base

- ✓ la requête de proximité est une version détendue de la requête par phrase et peut être considérée comme une **combinaison de termes et de phrases**.
- ✓ Les requêtes de proximité recherchent les termes de la requête à proximité les uns des autres.
- ✓ La proximité est utilisée comme facteur de classement des documents ou des pages retournés.

de données

## REQUÊTES DE PROXIMITÉ

- Par exemple, un document contenant tous les termes de requête proches d'eux-mêmes est considéré comme plus pertinent qu'une page dans laquelle les termes de requête sont éloignés.
- Certains systèmes permettent à l'utilisateur de spécifier la distance maximale autorisée entre les termes de la requête.
- La plupart des moteurs de recherche considèrent à la fois la proximité à long terme et l'ordre des termes dans la recherche.

# REQUÊTES DE DOCUMENTS COMPLETS

- L'utilisateur souhaite trouver d'autres documents similaires à ceux du document de requête.
- Certains moteurs de recherche (par exemple, Google) permettent à l'utilisateur d'émettre une telle requête en fournissant l'URL d'une page de requête.

# REQUÊTES DE DOCUMENTS COMPLETS



14

- En outre, dans les résultats renvoyés d'un moteur de recherche, chaque extrait peut avoir un lien appelé «plus comme ceci» ou «pages similaires».
- Lorsque l'utilisateur clique sur le lien, un ensemble de pages similaires à la page de l'extrait est renvoyé
-

# QUESTIONS EN LANGAGE NATUREL

- Le cas le plus complexe et l'idéal.
- L'utilisateur exprime son besoin d'information en tant que question en langage naturel. Le système trouve alors la réponse.
- Cependant, ces questions sont encore difficiles à gérer en raison de la difficulté de la compréhension du langage naturel.



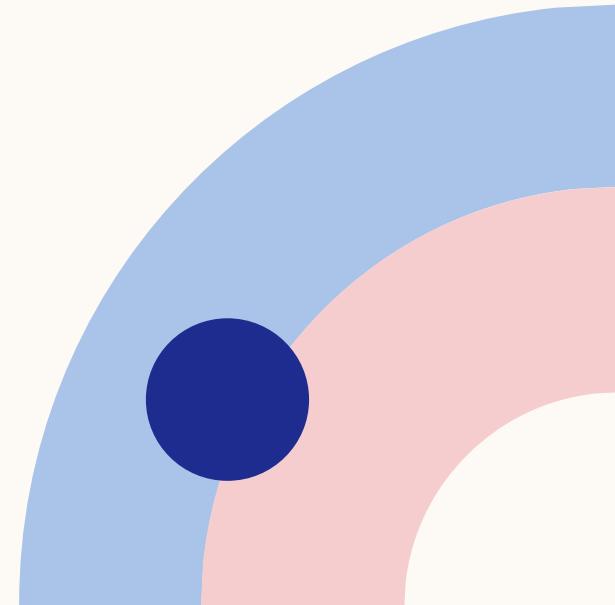
# QUESTIONS EN LANGAGE NATUREL

- Néanmoins, il s'agit d'un domaine de recherche actif, appelé question de réponse.
- Certains systèmes de recherche commencent à fournir des services de réponse aux questions sur certains types de questions spécifiques,

# QUESTIONS EN LANGAGE NATUREL

- Par exemple, questions de définition, qui demandent des définitions de termes techniques.
- Les questions de définition sont généralement plus faciles à répondre parce qu'il existe des modèles linguistiques forts indiquant des phrases de définition, par exemple, "défini comme", "se réfère à", etc.
- Les définitions peuvent généralement être extraites hors ligne,

# MODÈLES DE RI



**MERCI**