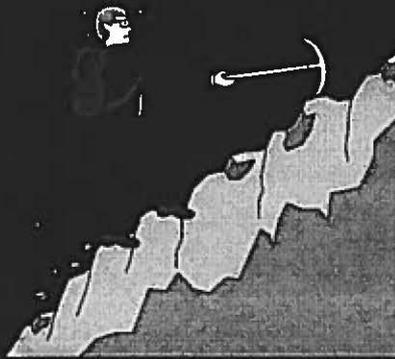


Recherche sur internet : Moteur...



**Chercher ou trouver
sur internet ?**

Laurence Jacqmin
Généraliste de Banque

1 Recherche sur internet : Moteur...

Chercher ou trouver sur internet ?

Laurence Jacqmin
Généraliste de Banque

2 Entre le mythe...

- Internet, c'est génial, tu tapes ce qui te passe par la tête et tu trouves toujours quelque chose d'intéressant !

3 ... et le cauchemar

- C'est l'horreur, tu as beau faire des requêtes complexes, tu passes des heures à dépouiller des milliers de réponses sans être sûr de trouver ce que tu cherches...

4 Le Web en chiffres

- Sites : 250.000 (FT, janvier 1997)
19.000 , août 1995
- Pages: 90 à 150 millions
- Accroissement : 170.000 pg/jour
- Moteurs de recherche : 47

5 pour faire face au développement exponentiel du web...

- 1992-93 : gopher
- outils de recherche sur gopher : Veronica,...
- moteurs de recherche sur le web lui-même

6 Définition

Un moteur de recherche contient:

- un robot
- un index
- un outil de recherche documentaire

7 typologie

8 Altavista

- Index généraliste
- 30 millions pg
- maj complète en une semaine
- zones indexées : pages web
- and not, or (implicite)
- troncature à gauche (*)
- Digital

9 Alta vista : Live topic

- affinement de la requête
- suggestion de termes associés
- réseau sémantique généré dynamique-ment à partir des textes

10 Lycos

- Index et répertoire généraliste
- 70 millions pg
- zones indexées : url, titres, header, 100 mots significatifs, 20 premières lignes
- and or not
- troncature implicite à droite
- Carnegie Mellon Univ., Center for MT

11 Infoseek

- Index généraliste
- 50 millions pg
- maj constante
- zones indexées: web, email, news
- and, or, not
- Infoseek Inc.

12 Infoseek Ultra

13 Yahoo

- Répertoire généraliste
- ? millions pg
- index hiérarchique à 5 niveaux
- zones indexées: titre, url, résumé
- and or
- Yahoo

14 Excite

- Répertoire généraliste
- 50 millions pg, 60.000 sites
- Index hiérarchique à 4 niveaux
- ou implicite
- Architext Software

15 Moteurs spécialisés

- Géographiques
 - B : Ad Valvas
 - F: Nomade
Francité
 - UE : Euroferret
Euroseek

16 Comparaison

17 Problèmes

- recherche sur index
-> différée
- recherche sur full text
- mécanismes d'exclusion
- text invisible
- achat de mots clés

18 Alternatives

- Moteurs
- agents (intelligents ?)
- webcasting et techniques "push"

19 Impact sur le métier

toujours un intermédiaire entre le client et l'information mais:

- Valideur (connaissance et évaluation des sources)
- formateur

20

Merci de
votre attention

laurence.jacqmin@gbank.be