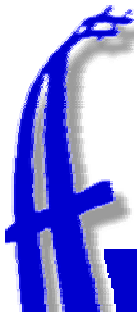




# Trouver de l'information sur le web

Jacques Monnard & Gérald Collaud  
Centre NTE, Université de Fribourg



Formation continue, 16 novembre 2000





# Programme

- Que peut-on trouver sur le web ?
- Qu'est-ce qu'un moteur de recherche ?
- Qu'est-ce qu'un index thématique ?
- Qu'est-ce qu'un méta-moteur ?
- Comment fonctionnent ces outils ?
- Quelles stratégies utiliser ?
- Comment améliorer mes recherches ?



# Objectifs

- Comprendre comment fonctionnent les moteurs de recherche
  - Savoir ce qu'il est possible de trouver
  - Savoir choisir entre un index thématique, un moteur de recherche, un méta-moteur, etc.
- Savoir chercher de l'information (pages web, email, logiciels, etc.)





# Plan

- Internet, web et types d'information
- Recherche de pages web
- Recherche d'autres types d'information
- Trucs et astuces
- L'avenir de la recherche sur le web



# Introduction

- Qu'est-ce qu'internet ?
- Qu'est-ce que le web ?
- Qu'est-ce qu'une ressource ?



# Ressources sur le web

http://

protocole  
utilisé

iiufpc01.unifr.ch

adresse de la machine

/nte/guides/

répertoire

default.html

nom de la  
ressource

#ancre

position dans  
le document



# Caractéristiques

- Internet : quelques chiffres
  - ◆ 300 mio d'utilisateurs
  - ◆ 1 mia de pages en décembre 1999, dont 85% en anglais
  - ◆ 7 mio. serveurs
- Recherche sur le web :
  - ◆ Indexation partielle :
    - taux de couverture d'environ 40%
    - Web invisible (bases de données, documents word, rtf, ...)
  - ◆ Instabilité
  - ◆ Renouvellement continu



# Quel format pour les informations ?

- Texte
  - ◆ .html, .htm
- Multimédia
  - ◆ Images : .gif, .jpeg
  - ◆ Sons, vidéos : .mp3, .mov, .rm, .mpg, ...
- Autres
  - ◆ documents formatés : .pdf, .doc, ...
  - ◆ logiciels : .exe, ...
  - ◆ fichiers compressés : .bin, .hqx, ...
  - ◆ ...





## 2. Recherche d'informations générales sur le web

- Les principaux outils à disposition:
  - ◆ Moteurs de recherche
  - ◆ Méta-moteurs de recherche
  - ◆ Index thématiques (annuaires)
  - ◆ Sites spécialisés



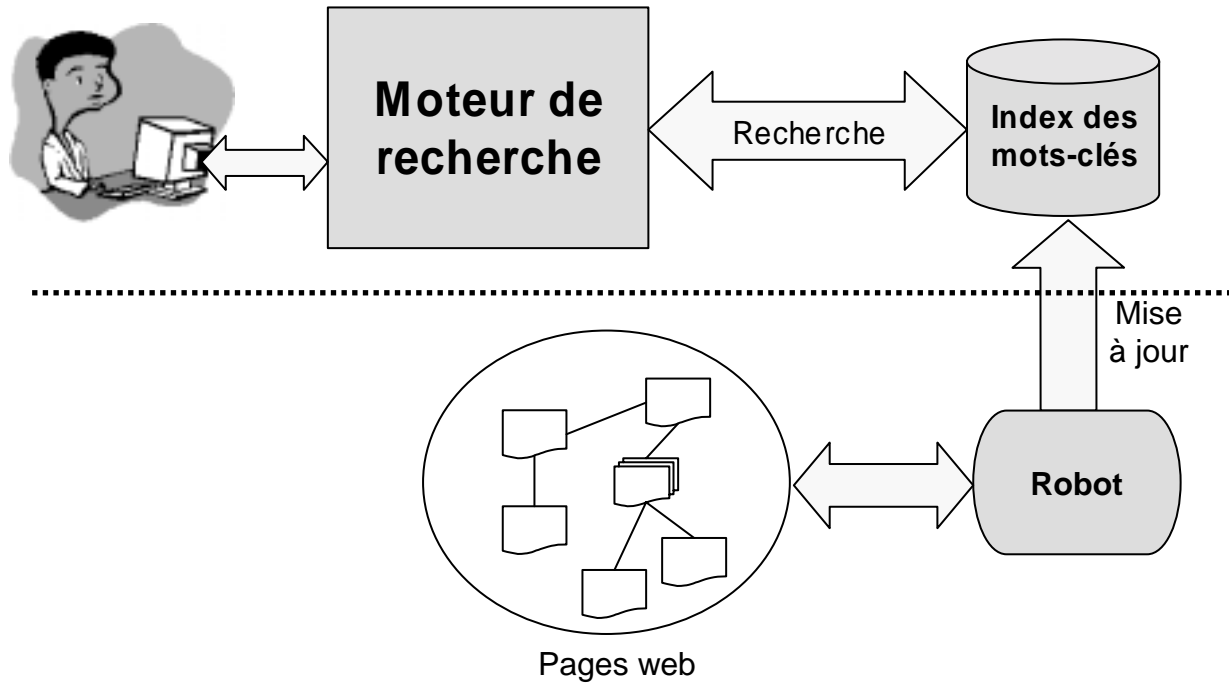


## 2.1 Stratégie de recherche

- Définir le type d'information qu'on cherche
- Sélectionner un outil de recherche:
  - ◆ recherche générale -> moteur de recherche ou méta-moteur
  - ◆ recherche dans un domaine spécifique -> index thématique ou site spécialisé
- Choisir des mots-clés appropriés et lancer la recherche
- Affiner la recherche si nécessaire



## 2.2 Moteurs de recherche: fonctionnement





# Moteurs de recherche

- Index créé par un «robot» qui parcourt régulièrement le web
- Avantages:
  - ◆ grande quantité de pages indexées
  - ◆ index local des mots-clés: rapidité et puissance de recherche
- Inconvénients:
  - ◆ aucune garantie de qualité
  - ◆ couverture limitée
  - ◆ mise à jour prend du temps
  - ◆ certains types de documents ne sont pas indexés (Word, PDF)
  - ◆ logique de plus en plus commerciale

# Quelques moteurs de recherche

- Francophones:
  - ◆ [AltaVista](#), [Lokace](#), [Voila](#)
- Suisse:
  - ◆ [Sear.ch](#), [Search.ch](#)
- Autres:
  - ◆ [Fast](#), [Google](#), [Go](#)

## 2.3 Les méta-moteurs de recherche

- Permettent d'effectuer une recherche simultanée dans plusieurs moteurs
- Résultats présentés sous forme intégrée
- Avantage: rapidité de recherche
- Inconvénient: peu adaptés pour des recherches complexes
- Exemples:
  - ◆ [MetaCrawler](#)
  - ◆ [Search.com](#)
  - ◆ [Dogpile](#)

## 2.4 Les index thématiques

- Adresses classées par catégories et sous-catégories
- Avantages:
  - ◆ sites sur un thème regroupés dans la même catégorie
  - ◆ qualité des informations
- Inconvénients:
  - ◆ taux de couverture du web faible
  - ◆ mise à jour lente
  - ◆ classement pas toujours évident
- Exemples:
  - ◆ Yahoo ([F](#), [D](#), [Monde](#)), [Francité](#), [Nomade](#), [Romandie.com](#)



## Faire apparaître un site dans un moteur/index

- Moteur de recherche: annonce préférable
- Index thématique: annonce nécessaire
- Fonction: “Annoncer un site”, “Ajouter une URL”
- Se limiter aux 10 principaux moteurs et index
- L’indexation peut prendre du temps!
- Utiliser les balises “meta” *keyword* et *description*







## 2.5 Les sites spécialisés

- Centres de ressources sur un domaine spécifique
- Réalisés par des spécialistes d'un domaine
- Taille plus petite que les annuaires commerciaux
- Haute valeur ajoutée



# Comment trouver un site spécialisé

- Recherche (p. ex. Avec Yahoo)
- Regroupement de sites spécialisés :
  - ◆ [About](#), [WWW Virtual Library](#)
- Annuaire de sites spécialisés:
  - ◆ [Signets de la BNF](#), [The Argus Clearinghouse](#), [TheBigHub](#)
- Existent dans de nombreux domaines:
  - ◆ enseignement : [Educa](#), [EducaSource](#)
  - ◆ création de sites web : [AaZ Webmasters](#)

# Comment chercher quoi ?

- Pages web et images
- Personnes (adresses électroniques)
  - ◆ sites dédiés
  - ◆ site de l'entreprise
- Logiciels
- Forums de discussion
  - ◆ Catégories principales
  - ◆ Sites
- «Ask an expert»
- Bases de données

## 3.1 Recherche de personnes

- Sites dédiés:
  - ◆ [Four11](#)
  - ◆ [WhoWhere](#)
  - ◆ [MESA](#)
- Aller sur le site de l'entreprise/institution dans laquelle travaille la personne.





## 3.2 Recherche de logiciels

- Logithèque Yahoo
- CNET Shareware
- Download.com
- ZDNet
- Tucows



## 3.3 Recherche de discussions

- Différentes catégories de discussions (newsgroup):
  - ◆ alt, misc, bionet, comp, sci, talk, news, soc, ...
  - ◆ par pays: ch.talk, fr.culture, de.markt,...
- Pour la recherche:
  - ◆ [Voila](#) (francophone)
  - ◆ [Deja](#)



## 3.6 Ask an expert

- Pour poser une question à un expert (humain):
  - ◆ [AllExperts](#)
  - ◆ [Experts Yahoo](#)
  - ◆ Ask [An Expert](#)
- Pour poser une question à un expert (automatisé):
  - ◆ [AskJeeves](#)



## 3.7 Web invisible - bases de données

- Les bases de données ne sont généralement pas indexées par les moteurs de recherche
- De nombreuses bases de données “intéressantes” sont payantes
- Sites recensant des bases de données :
  - ◆ <http://urfist.univ-lyon1.fr/gratuits/index.html> (BD gratuites)
  - ◆ The Invisible Web
  - ◆ CompletePlanet
- Meta-recherche de bases de données: [Search.com](http://www.search.com)
- Journaux: [PresseWeb](http://www.presseweb.com)



## 4. Trucs et astuces

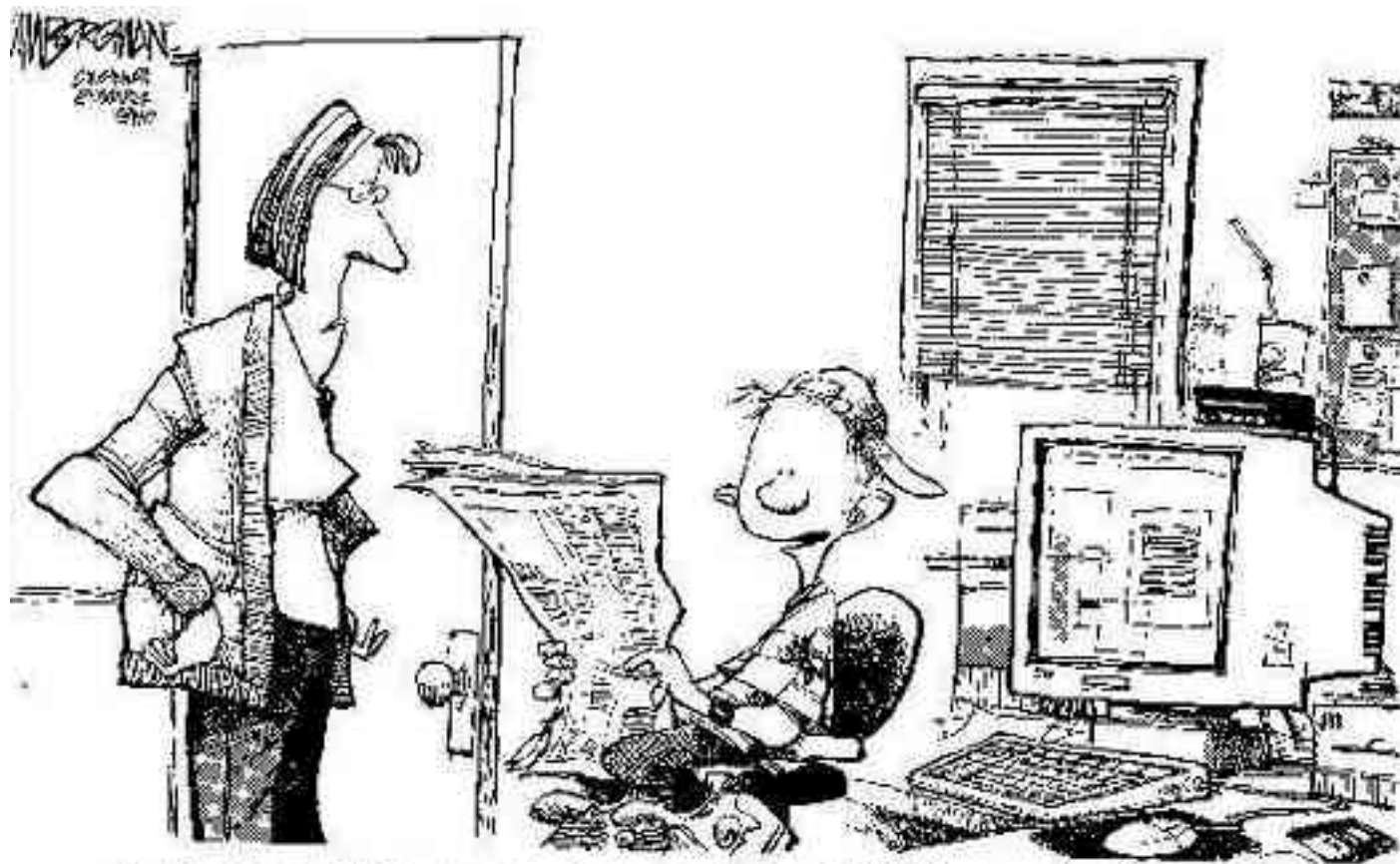
- Bon sens : [www.fribourg.ch](http://www.fribourg.ch), [www.unifr.ch](http://www.unifr.ch), ...
- Effacer des parties de l'URL pour trouver d'autres parties du site
- Utiliser l'aide en ligne du moteur de recherche
- Les signets sont souvent une bonne source d'information
- Amusant :
  - ◆ Sites qui ont créé un lien vers le mien (avec Altavista)
  - ◆ Ce que les gens cherchent - Top ten searches:  
<http://www.excite.com/search/voyeur/>



# L'avenir de la recherche sur le web...

- Logiciels installés localement:
  - ◆ Sherlock (Mac)
  - ◆ Copernic (Windows)
- Gnutella: recherche décentralisée
- **Tout change tout le temps!**





Intéressant... C'est comme un fichier portable de 500 Ko et on n'a même pas à attendre la fin du téléchargement... Et vous dites que ça s'appelle "un journal" ???