

QUAND LES SURFEURS SE NOIENT : SOYEZ CHERCHEURS

Cette séance va être l'occasion d'expérimenter les diverses grandes techniques de recherche d'informations sur Internet. Dans votre vie d'étudiant et ensuite, professionnelle, vous allez sûrement avoir à constituer des dossiers, des états de l'art, faire de la bibliographie ... Internet est un outil formidable mais uniquement pour celui qui sait le parcourir efficacement (ce que nous avons effleuré lors de la dernière séance), y chercher et y trouver l'information.

PRELIMINAIRES

Lancez Netscape.

Pour répondre aux questions suivantes, il vous faudra exploiter le document sur les techniques de recherche d'informations sur Internet. Toutes les manipulations que l'on vous demande de faire sont expliquées dans cette fiche. Les réponses sont à reporter sur la feuille distribuée à cet effet et à rendre en fin de séance.

Une recherche s'initie à partir de la page d'accueil d'un outil de recherche, page accessible par son adresse web (URL). Parmi vos signets, ayez les URL de vos outils de recherche préférés.

QUESTION I : ANNUAIRES SIMPLES D'UTILISATION

- (a) Connectez-vous sur www.utilnet.fr
- (b) Dans une autre fenêtre de navigation, connectez-vous sur un autre annuaire, www.rivieraworld.com
- (c) Parmi ces 2, sauriez-vous dire lequel est thématique ? Répondre sur la feuille jointe
- (d) Dans utilnet, où se trouvent les sites populaires ? Répondre sur la feuille jointe.

Vous avez besoin de vacances, et vous êtes un(e) fêtarde(e). De plus, vous adorez les festivals.

- (e) Sous utilnet, affinez vos sujets, afin de trouver l'URL vous permettant de choisir votre destination.
- (f) Sur la feuille jointe, donnez le parcours par sujets de plus en plus précis que vous avez suivi, et l'URL à laquelle vous êtes finalement arrivé.

QUESTION II : ANNUAIRE + MOTEUR

Connectez-vous sur le site français de Yahoo! dont l'URL est fr.yahoo.com. Commençons par expérimenter l'aspect purement annuaire de Yahoo !

Ce soir, vous aimeriez bien lire une petite histoire de Science fiction, mais il est déjà trop tard, la librairie en face de chez vous est fermée. Si vous alliez chercher un livre dont le texte est en ligne ?

- (a) A partir de la liste des sujets principaux, trouver votre chemin jusqu'à une liste de livres de Science fiction dont les textes sont en ligne.
- (b) Donnez le chemin parcouru sur la feuille jointe, ainsi que le titre du premier de ces recueils.

En fait dans cette liste, vous voyez apparaître l'auteur Stephen King, et cela vous rappelle que l'autre jour, vous vous étiez demandé s'il avait récemment publié un roman, et si oui quel est son titre. Nous allons donc utiliser l'aspect Moteur de recherche de Yahoo !

- (c) Commencez par rechercher les références à Stephen King, répertoriés dans Yahoo !
- (d) Cet auteur est-il si célèbre qu'il a sa propre catégorie dans Yahoo ! ? Répondez sur la feuille jointe.
- (e) Y-a-t-il un site privilégié où vous allez savoir le titre de son dernier roman ? Donnez l'URL sur la feuille jointe.
- (f) Supposons que vous ne trouviez pas votre bonheur, alors continuez la recherche sur Stephen King à l'aide d'un moteur de recherche extérieur à Yahoo!, en

cliquant "Les 3 résultats suivants" puis "Autres Pages Web". Quel est le nom de ce moteur ?

- (g) Indiquez sur la feuille jointe combien de réponses sont obtenues.
- (h) Y en a-t-il parmi elles où seulement le mot "king" apparaît ? Si oui, refaites la recherche en considérant Stephen King comme un seul mot-clé et répondez sur la feuille jointe, en donnant aussi le nombre de réponses ainsi obtenues.

QUESTION III : RECHERCHE SUR UN SEUL MOT, A L'AIDE D'UN MOTEUR

Nous allons expérimenter la recherche par mot clé, en nous servant du puissant moteur Altavista. On utilisera l'URL www.altavista.com, mais il existe son équivalent français à www.altavista.fr

Chargez dans 2 fenêtres de navigation les pages d'accueil de ces 2 sites.

- (a) Constatez la différence en terme de publicité et de nouvelles
- (b) Constatez aussi que ce moteur tend à être un annuaire. Pourquoi ? (répondez sur la feuille jointe).

Supposons que vous cherchiez à joindre toutes les personnes du monde, qui dans leur cursus universitaire ont suivi, ne serait-ce qu'en partie, la filière MASS. En effet, vous voulez monter une association des anciens de Mass, et cherchez donc à contacter le plus de personnes possibles, via Internet. Notez que l'on ne cherche pas des informations sur la filière MASS en elle-même, donc il faudra à terme exclure de votre résultat de recherche, les pages de présentation de ces filières. Il s'agira plutôt

d'essayer d'obtenir dans la liste de résultats des pages correspondant à des curriculum vitae. Comme certaines personnes se présentent autrement qu'en français (par exemple, en anglais), nous ne nous restreindrons pas aux pages écrites en français. Finalement, comme les pages de ces personnes ne sont pas forcément sur des machines du domaine Internet .fr, nous ferons une recherche sur tout le Web.

- (c) Lancez votre requête ainsi : mass
- (d) Indiquez le nombre (approximatif) de réponses sur la feuille jointe. Indiquez un des thèmes obtenus, qui n'a rien à voir avec la filière MASS.
- (e) Lancez votre requête ainsi : MASS, puisque c'est un acronyme.
- (f) Indiquez le nombre de réponses sur la feuille jointe.
- (g) Juste pour voir, demandez une recherche en restreignant aux pages écrites en français. Dans la première page de réponses, vous avez celle-ci "ARISING ATTITUDE - MANIFESTO WEBSITE" ? Pourquoi ? Pour y répondre, allez sur ce site, et en demandant le source de la page, repérez où apparaît MASS. Répondre sur la feuille jointe.
- (h) Vous aviez pourtant demandé les pages écrites en français, et celle-ci semble plutôt écrite en anglais. Mais en y regardant de plus près, des mots français apparaissent. Notez-en un sur la feuille jointe.

QUESTION IV : RECHERCHE SUR PLUSIEURS MOTS

MASS semble trop vague (puisque nous avons beaucoup trop de réponses), nous allons donc orienter notre recherche à l'aide des termes mathématiques, appliquées, sciences, sociales.

- (a) Lancez votre recherche en écrivant ces 4 termes en minuscules, sans les relier par des guillemets et sans accents.
- (b) Au vu des premiers résultats obtenus, avons-nous bien fait de formuler notre requête en minuscules, sans accents ?
- (c) Parmi les milliers de réponses obtenues, sommes-nous sûrs que les 4 termes apparaissent ? Forçons donc leur

présence dans la page et indiquez sur la feuille jointe de quel pourcentage environ la quantité de réponses a diminué.

- (d) Tout le monde met-il un "s" à mathématiques ? Ce serait dommage de rater le CV de quelqu'un pour cette unique raison. Reformuler votre requête et l'indiquer sur la feuille.
- (e) Mais au fait, dans les quelques milliers de pages qu'il nous reste, rien ne nous dit que les termes 'mathématique' et 'appliqué' sont en rapport ? Reformuler votre requête pour que ce soit le cas, en pensant à faire de même pour sciences et sociales. Indiquer sa forme sur la feuille jointe.
- (f) Avez-vous remarqué que certaines formulations sont "mathématiques appliquées AUX sciences sociales", alors que d'autres sont "mathématiques appliquées ET sciences sociales" ? Quel est le résultat si nous formulons notre requête ainsi : "mathematique* applique* sciences sociales" ? Répondez sur la feuille jointe.
- (g) Quel est le résultat si nous formulons notre requête en mettant une * pour remplacer aussi bien "et" que "aux" que "&" dans la phrase ? Donnez le nombre de réponses sur la feuille jointe après avoir vérifié sur différentes pages obtenues, que l'* a bien remplacé les termes "et", ou "aux".

QUESTION V : RECHERCHE AVANCEE

Maintenant que nous avons fait le tour des subtilités autour de MASS et des 4 termes qui se cachent sous cet acronyme, nous allons poursuivre notre recherche en partant du principe que si dans un CV on écrit que l'on a suivi la filière MASS, on va probablement, pas bien loin, décrire ce que veulent dire ces 4 lettres.

L'opérateur NEAR (proximité dans le texte) va nous être bien utile.

- (a) Passez en mode recherche avancée.
- (b) Exprimez votre requête de telle sorte que ces 4 termes et MASS soient proches. Indiquez la formulation sur la feuille jointe.

(c) Cet opérateur commute-t-il ?

Avez-vous remarqué que de nombreuses pages proposées comme résultat, décrivant une filière MASS, comprennent ce sigle ou sa description, dans la page de titre. Nous allons compter combien il existe de telles pages, en préfixant par **title**: l'acronyme MASS d'une part, et en faisant de même pour la phrase contenant les 4 termes. Ces 2 parties de notre formulation seront donc séparées par un OR, puisque dans la page de titre, il n'est pas obligatoire qu'apparaissent l'acronyme et sa description.

(d) Donnez la formule obtenue sur la feuille jointe.

Le nombre de réponses est à nouveau très important, mais tout va s'arranger, puisque nous voulons exclure ces pages.

(e) Donnez une formule (qui commence à être compliquée !) nous permettant de sélectionner les pages qui contiennent "mathematique* applique* * sciences sociales" NEAR MASS, mais qui excluent les pages dont c'est le titre. Indiquez votre formule sur la feuille jointe.

On a encore 242 pages ! Essayons de ne garder que celles qui sont des curriculum.

(f) Tapez cette formule :

```
((("mathematique* applique* * sciences sociales" NEAR MASS) AND (CV OR curriculum)) AND NOT (title:MASS OR title:"mathematique* applique* * sciences sociales" )
```

et donnez le nombre de pages obtenues. Il ne vous reste plus qu'à contacter ces quelques personnes, sachant que à cause de la contrainte due au mot CV, nous sommes peut-être passés quand même à côté de certaines répondant à nos critères. Mais c'est la limite de la recherche par mot-clé.