

RIM, laboratoire no 1

Matière	Recherche d'informations multimedia
Professeur	Nastaran Fatemi
Assistant	Florian Poulin
Auteurs	Simon Hintermann, Romain de Wolff

I. Introduction

Ce premier laboratoire porte sur l'étude du fonctionnement au niveau utilisateur et au niveau interne des requêtes sur les moteurs de recherche.

II. Objectifs

Notre choix s'est porté sur le moteur de recherche « Yahoo.com », que nous n'utilisons d'ailleurs pas d'ordinaire.

• Première partie

Il s'agit de définir les options de recherche mises à disposition par le moteur de recherche.

• Deuxième partie

Le but est de créer une requête pour trouver un bon restaurant. On va ensuite, en fonction des résultats obtenus, essayer d'analyser la fonction de correspondance ainsi que la méthode de pondération qu'utilise « Yahoo.com ».

• Troisième partie

Description des éléments mis à disposition dans les résultats d'une recherche.

III. Réponses aux questions

• Première partie

En arrivant sur Yahoo, nous avons à disposition une recherche par mots-clé, mais nous pouvons aller dans la recherche avancée, qui nous ouvre des possibilités d'affinage intéressantes.

On peut affiner l'affichage des requêtes avec quatre boxes permettant respectivement de:

- Afficher les résultats comprenant tous les mots spécifiés

YAHOO! SEARCH

☒ the Web ☐ just pages from Switzerland

Search Results

1. [Outils mi-hauteur Sneeboer - Jardins Animes](#) - [Translate this page](#)
- Outils de **Jardin** - Outils mi-hauteur Sneeboer chez ... Décoration de **jardin** Animaux en bronz
Animaux cracheurs en ... 05 **Pelle** rateau Sneeboer. Nos 20 ...
fr.jardins-animes.com/outils-jardin-outils-mihauteur-sneeboer-c-4_6... - 34k - [Cached](#)
2. [Outils à une main Sneeboer - Jardins Animes](#) - [Translate this page](#)
- Outils de **Jardin** - Outils à une main Sneeboer chez Jardins Animes ... **Pelle** à désherber
longue Sneeboer. 32,30. € Grattoir scratch Sneeboer. 32,30 ...
fr.jardins-animes.com/outils-jardin-outils-une-main-sneeboer-c-4_7... - 47k - [Cached](#)
3. [Fiskars® - Jardin](#) - [Translate this page](#)
Pelle de secours ultra légère. **Pelle** manche bois - tête en acier forgé. **Pelle** pliable avec étui
coton © 2006 FISKARS BRANDS, INC. ...
fiskars.com/FR/Jardin/.../Pelle/ProductListing1.html - 11k - [Cached](#)

- Afficher les résultats contenant une phrase exacte

YAHOO! SEARCH

☒ the Web ☐ just pages from Switzerland

Search Results

1. [LE FORUM DU CHIEN !](#) - [Translate this page](#)
Il n'est ni un enfant (même s'il comble un vide) , ni un conjoint ... ne "débordent" pas , il n'en ser
peut etre pas de même pour **le chien du voisin**. ...
groups.msn.com/LEFORUMDUCHIEN/liste.msnw - 20k - [Cached](#)
2. [Yahoo! Questions/Réponses - "si le fils,le neveu,la femme et le chien ...](#) - [Translat](#)
[this page](#)
5 answers - plusieurs fois entendu dans différentes émissions ... le neveu,la femme et **le chien**
du voisin" ont retrouvé un travail Sarkozy aura réussi? ...
qc.answers.yahoo.com/question/index?qid=20070921035603AAJ1KTB&show=7 - 86k - [Cache](#)
3. [Comment tuer le chien de son voisin](#) - [Translate this page](#)
Mais ce qui l'insupporte le plus au monde, c'est **le chien du voisin** qui ne cesse d'aboyer... A
Los Angeles, Peter McGowan, un écrivain frustré, traverse ...
www.allocine.fr/film/fichefilm_gen_cfilm=33375.html - 156k - [Cached](#)

- Afficher les résultats comprenant au moins un des mots spécifiés

YAHOO! SEARCH

☒ the Web ☐ just pages from Switzerland

Search Results

1. [pelle pelle](#)
Please install Flash and turn on Javascript. ...
www.pellepelle.com - 4k - [Cached](#)
2. [Flickr: Photos tagged with jardin](#)
Flickr is almost certainly the best online photo management and sharing ... **Jardin**. Electric tea
maker automatically adjusts for different tea. www. ...
www.flickr.com/photos/tags/jardin - 34k - [Cached](#)
3. [Jardin - Wikipedia, la enciclopedia libre](#) - [Translate this page](#)
Jardin ... Un **jardin** puede incorporar tanto materiales naturales como ... 1 Tipos de **jardin**. 2
Historia. 3 Los jardines en la literatura. 4 Véase también ...
Más sobre el tema: [Tipos de jardín](#) - [Historia](#) - [Los jardines en la literatura](#)
es.wikipedia.org/wiki/Jardin - 26k - [Cached](#)

- Afficher les résultats ne comprenant aucun des mots spécifiés



Ces quatre boxes comprennent aussi à la fin une option permettant de sélectionner l'occurrence des mots spécifiés dans la page, le titre, ou l'URL.



Illustration 1: Recherche avec le mot "jardin dans l'URL"

Ces quatre options peuvent être aussi utilisées sous la forme d'opérateurs dans la recherche standard de Yahoo. Ces opérateurs sont expliqués sur Yahoo (1). Nous avons à disposition:

(1) <http://eur.help.yahoo.com/help/fr/ysearch/tips/tips-04.html>

- Les guillemets permettent de chercher une phrase exacte
- Le signe + devant un mot permet de forcer la recherche à lister des sites contenant impérativement ce mot
- Le signe – devant un mot permet de forcer la recherche à lister des site ne contenant pas ce mot

- L'opérateur 'OR' permet de faire un ou logique entre deux mots-clé
- Les mots-clé de recherche sont automatiquement liés par l'opérateur AND, dans une recherche standard.

Nous pouvons encore spécifier la date de mise à jour des pages recherchées.



Illustration 2: Recherche d'une site contenant 'jardin' mis à jour lors des 3 derniers mois

Le domaine de la recherche peut être imposé (.com, .ch, etc...), ainsi que le site concerné, si on veut aller encore plus loin (lematin.ch, p. ex.).



Illustration 3: Recherche avec 'jardin' limitée au site lematin.ch

Le format du fichier recherché peut être précisé (HTML, PDF, Excel, Word, PowerPoint, RSS/XML, Texte).

1. flarere.free.fr/Autres/Francais/2002_10/Ma_cabane_au_fond_du_jardin.txt
D A D D G ma Cabane au fond du JarDin moi j'y Vais quand j'ai beSoin G G Bm ... petit Coin ma cabane au fond du jarDin Elle trône tout au bout d'un chemin plein ...
flarere.free.fr/Autres/Francais/.../Ma_cabane_au_fond_du_jardin.txt - 3k - [En cache](#)
2. www.figuresdestyle.com/fyfield/jardin/WS_FTP.LOG
... figuresdestyle/rose/jardin interview2.htm 2000.09.20 13:53 B L:\Figures de style ...
figuresdestyle/mjrose/jardin sites.html 2000.09.26 16:24 B L:\Figures ...
www.figuresdestyle.com/fyfield/jardin/WS_FTP.LOG - 2k - [En cache](#)
3. www.taupedelire.com/new/evolutions.txt
... du texte de l'événement lorsqu'une taupe est crée dans le jardin d'enfant. ... Une sortie du jardin d'enfants par déplacement ou une orientation de la 7b vers ...
www.taupedelire.com/new/evolutions.txt - 62k - [En cache](#)

Illustration 4: Recherche de 'jardin' en fichier texte

Le pays sur lequel appliquer la recherche peut être précisé.

1. [Le Jardin](#)
CONTACT. Le Jardin Boutique Villa, Jl. Sarinande 7, Seminyak, Kuta - Bali - 80361 ... le jardin.
room type, service & facilities, reservation, packages, contact ...
www.lejardinvilla.com/boutique-villa-contact.html - 6k - [En cache](#)
2. [brosurjawa](#) (PDF)
jardin botanique mondialement réputé de Bogor. ... Fondé en 1817, ce jardin est un "musée" de milliers ... jardin est également utilisé à des fins de recherche. ...
www.indonesia-mrs.com/brochures/brosurjawa.pdf - 87k - [Afficher en html](#)
3. [LES MOLUQUES](#) (PDF)
<http://www.indonesia-mrs.com>. INDONESIE. LES MOLUQUES. JARDIN DE CORAIL SOUS MARINS. A Pombo Island (île de Pombo) ne pas loin des plages, des récifs ...
www.indonesia-mrs.com/brochures/brosmaluku.pdf - 95k - [Afficher en html](#)

Illustration 5: Recherche de 'jardin' limité à l'Indonésie

On peut aussi préciser la langue dans laquelle on veut avoir les documents.

1. [Hotel Restaurant Jardin, Bern](#) - [Traduire cette page](#)
Ein Stadthotel mit der familiären Note. Das Hotel liegt fernab vom Verkehrslärm, mit Blick ins Grüne und doch nahe vom Zentrum.
www.hotel-jardin.ch - 5k
2. [Jardin Tecina • Hotel Jardin Tecina La Gomera • JARDIN-TECINA.DE](#) - [Traduire cette page](#)
Hotel Jardin Tecina zum Bestpreis buchen mit Preisvergleich aller Veranstalter ... Hotel Jardin Tecina • Highlights • Fotos • Buchen • Gomera • Playa Santiago ...
www.jardin-tecina.de - 13k - [En cache](#)
3. [Jardin de Palme - Apartments](#) - [Traduire cette page](#)
www.jardin-de-palme.com ... Um sie zu betreten, passieren Sie unsere Eingangshalle von der "Avenue San Brancat" aus. Hier finden Sie originale und frei ...
www.jardin-de-palme.com/de/appartements.php - 19k - [En cache](#)

Illustration 6: Recherche de 'jardin' en allemand

(le moteur de recherche ne traduit pas 'jardin' en allemand...)

Finalement, un filtre du contenu pornographique peut être appliqué (mais sans garantie).

Il est à noter que les option avancées de recherche peuvent faire l'objet d'instructions

dédiées lors de la recherche par mots-clé. Par exemple, rechercher des pages appartenant au site heig-vd.ch sera un « paramètre », traduit par *site:heig-vd.ch* dans la ligne saisie des mots-clé.

• Deuxième partie

Pour formuler une requête nous permettant de trouver un bon restaurant, nous devons déjà préciser dans quelle ville nous voulons rechercher, car la recherche de Yahoo ne prend pas en compte la position géographique de l'utilisateur.

Le fait de rechercher des restaurants en suisse, avec les options avancées ne fonctionne d'ailleurs pas bien (du tout), car il nous sort quand même des sites représentant des restaurants en France en tête de liste.

Le fait de chercher un « bon » restaurant ne pourra pas être précisé dans les mots-clé, bien entendu, mais si les résultats sont classés suivant leur référencement, on pourra en tirer que ce sont des restaurants connus, donc appréciés.

En l'occurrence, sur Yahoo, nous avons tapé la requête 'restaurant lausanne', pour n'avoir que des restaurants de la région. Les résultats contiennent bien tous les deux mots-clé, et concernent des restaurants, mais on ne peut pas dire avec certitude que les premiers résultats soient les meilleurs. Quand on voit ce résultat en 3ème position, par exemple:

3. [Restaurant Romanel - Cafe Brasserie Vaud - Kebab Lausanne - Traiteur ...](#)
... Centre commercial Migros de Romanel sur **Lausanne**, La Chope d'Or est fière de ... entrez ici
© Copyright CsetID 2007 Création de site internet - plan de site ...
www.restaurant-brasserie-suisse.com/index.htm - 5k - [En cache](#)

on peut penser que la pondération prend en compte le référencement des sites.

La fonction de correspondance à l'air de chercher simplement à trouver des pages dans lesquelles les mots-clé apparaissent (dans le titre surtout).

Une manière plus intelligente de trouver un « bon » restaurant est de taper 'restaurant classement lausanne', qui va nous donner des sites de classements de restaurants, nous permettant de voir l'avis d'autres personnes, avoir une idée un peu plus précise de l'endroit où l'on va aller manger...

1. [Resto-ranG - Guide interactif des restaurants - Lausanne](#)
... Afin de rendre le **classement** plus dynamique et inviter les ... 5 derniers restos ajoutés à **Lausanne**. Kai Zen restaurant. 7.70. Moulin d'Assens. 10.00 ...
www.resto-rang.ch/index.cfm?canton=LS - 57k - [En cache](#)

• Troisième partie

Lors d'une recherche, Yahoo va nous afficher plusieurs éléments pour donner des informations cohérentes en rapport avec les pages qu'il a trouvées.

1. [Jardin - Wikipédia](#)

Les parties du **jardin** : pelouse - étang ornemental - parterre - potager ... à un palais ou un château comme le **jardin** à l'italienne des villas toscanes. ...

Liens directs: [Les types de jardin](#) - [Les éléments du jardin](#) - [Les bâtiments du jardin ou le jardin dans le bâtiment](#)

fr.wikipedia.org/wiki/Jardin - 36k - [En cache](#)

Illustration 7: Simple recherche 'jardin'

- Le titre: il est repris de la balise <title> de la page, libre à chaque possesseur de site de mettre donc un titre explicatif. Ce titre sert à donner le sujet principal de la page.
 - Le texte contenu en-dessous sélectionne soit des morceaux de texte contenant le mot-clé (ou les mots-clé), soit une partie du texte de la page, si le mot-clé n'apparaît que dans le titre ou dans l'adresse.
 - (*) Les liens directs sont probablement disponibles avec l'inscription au moteur de recherche, ou encore à l'aide du code de la page... Ces liens sont des liens pertinents renvoyant à d'autres sujets. Ici encore, ils sont précisés par le propriétaire du site.
 - L'adresse du lien est affichée en vert. Elle est tronquée si trop grande.
 - La taille de la page est affichée en gris/bleu clair
 - Un liens sur la page en cache. Utile si la page n'est plus accessible sur le site « original ». Cela permet de quand même accéder à l'information.
- (*) Les liens directs apparaissent rarement. Ils sont affiché quand il existe des sous-rubrique dans le site avec les mots clés. Ce qui permet d'y accéder rapidement. Un exemple typique est Wikipedia, comme montré dans l'illustration ci-dessus.

• Quatrième partie

Kartoo.com: Meta-moteur de recherche permettant de rechercher des sites par sujets en créant à chaque sujet recherché une « carte » permettant de raffiner la recherche. L'utilisabilité du système est certes limitée, car le système de classement n'est pas créé par l'homme, mais calculé, donc on trouvera des mots-clé trop vagues ou n'évoquant rien pour nous, mais cela peut quand même être intéressant de voir les sites proposés évoluer au fil des mots-clé rajoutés.

Particularité: les résultats de recherche sont donnés sous forme graphique, et non de liste.

Ujiko.com: Meta-moteur de recherche plutôt basé sur l'aspect ludique de la recherche. On entre un mot-clé, et on obtient des point à chaque fois que l'on va visiter un site depuis ujiko. Chaque niveau passé à l'aide des point gagnés permet d'avoir une option de personnalisation de l'interface en plus. La recherche fonctionne à peu près comme Kartoo, on obtient des mots-clé représentant les sujets principaux ressortant de la recherche.

Particularité: les résultats(12) sont affichés autour des mots-clés, et passer sur chaque mot-clé fait apparaître ceux qui sont concernés par le mot-clé.

Foxxx.com: Moteur de recherche qui nécessite une inscription pour une utilisation de toute ses fonctionnalités. Il se définit comme outil de recherche collaboratif. Les recherches de tests effectuées donnent des résultats très approximatifs.

Particularité: les recherches sont effectuées sur plusieurs site de recherche (google, alltheweb,...)

Clusty.com: Site de recherche similaire à google. Il permet aussi de prévisualiser les sites et les ouvrir dans une nouvelle fenêtre. Des gadgets...

Particularité: permet de regrouper les recherche par « cluster ». Possibilité d'afficher la définition du/des termes recherchés.

Boolgum.fr: Site de recherche spécialisé pour la France. Il est aussi orienté vers une recherche pratique et ludique. Il permet de comparer des prix, rechercher, lister les synonymes, effectuer des recherches connexes. Il permet aussi de rechercher par code postal (pour la France toujours, un code postal suisse ne fonctionnant pas).

Particularité: En double cliquant sur un mot dans la liste des résultats, on obtient la définition dudit mot.

AllTheWeb.com: Moteur de recherche paramétrable mais sans grande particularité.

Ask: Permet de faire une recherche dans 5 catégories différentes, à savoir *Web*, *Images*, *City*, *News*, *Blogs*. Les recherches sont suggérées à la volée (AJAX). Les résultats de la recherche sont affichée sur une page à droite, similaire à Google. Sur la partie gauche

ce situe un panneau intéressant qui permet d'affiner ou d'élargir le champs de ses recherches avec des suggestions. La partie city est intéressante mais semble ne pas totalement fonctionner correctement (nous avons entouré la suisse et recherché « Lausanne », sans succès).

Songtapper: Site original qui permet de taper le rythme d'une chanson pour la retrouver. Pas très précis, mais pour des rythmes simples la recherche fonctionne bien.

Answers.com: Site censé répondre aux question, mais semble ne pas encore être très au point. Il est difficile d'avoir des réponses à ses propres questions. Toutefois, et heureusement, les exemples fonctionnent bien.

Wisenut.com: Site de recherche qui n'est vraiment pas très utile ni fonctionnel. Censé recherché *exactement* nous le trouvons flou et mal organisé, donc pas pratique à utiliser.

<http://labs.systemone.at/retrievr> : Site de recherches d'images que l'on fait en dessinant le contraste principal. Très approximatif, il ne semble pas du tout au point.

Local.ch: Site de recherche pour la Suisse qui permet de faire des recherches d'annonces, de loisirs, d'annuaires. Les finitions du site sont très bien faites. Il est agréable et facile à utiliser.

Particularité: affiche une carte qui montre où se trouve l'objet en question.

Krugle.com: Site de recherche pour développeur. Les recherches sont conservées dans des tabulations, des projets libres sont affichés sur la gauche. Grande source d'information très utile.

Particularité: Les catégories sont très bien faites et permettent de facilement naviguer dans le site.

IV. Conclusion

Ce laboratoire nous a surtout permis de découvrir d'autres moteurs de recherche dont nous ne connaissions pas l'existence. Les méta-moteurs sont les plus intéressants, alors que celui pour retrouver des codes-sources s'avère être très intéressant pour gagner du temps dans nos projets personnels.

A l'utilisation, Yahoo s'est avéré moins ergonomique que Google, qui reste une référence en matière de simplicité et d'efficacité.