



## **Domine D4: Organiser la recherche d'informations à l'ère du numérique**

Chafik Aloulou, Mohamed Koutheir khribi, Walid chaou, Nizar Kayati, Lilia Cheniti, Sonia Guanouchi, Houda Houissa

## Domaine D4 : Organiser la recherche d'informations à l'ère du numérique



- D4.1 : Rechercher de l'information avec une démarche adaptée
- D4.2 : Evaluer les résultats d'une recherche
- D4.3 : Récupérer et référencer une ressource numérique en ligne
- D4.4 : Organiser une veille informationnelle



جامعة تونس الافتراضية  
UNIVERSITE VIRTUELLE DE TUNIS

CERTIFICAT INFORMATIQUE & INTERNET

## **D4-1 : Rechercher de l'information avec une démarche adaptée**

Préparé par : Mohamed Koutheaïr KHRIBI

# Plan

I-Recherche d'information

II-Sources d'information

III-Outils de recherche d'information

IV-Formulation de requête

V- Récupération d'information

# I-Recherche d'information

## →Définition :

La recherche d'information RI (IR Information Retrieval en anglais) est la science qui consiste à **rechercher l'information dans des documents** (les documents eux-mêmes ou les métadonnées qui décrivent les documents), dans **des bases de données** , dans **le World Wide Web, l'internet, les intranets**, pour le texte, le son, les images, les données. Dans le jargon de la documentique, la recherche d'information est définie comme étant (Paris, ADBS, 2004) : « **l'ensemble des méthodes, procédures et techniques permettant, en fonction de critères de recherche propres à l'utilisateur, de sélectionner l'information dans un ou plusieurs fonds de documents plus ou moins structurés** ».

## →Définition (suite) :

La recherche d'information est un domaine historiquement lié aux sciences de l'information et à la bibliothéconomie qui ont toujours eu le souci d'établir des représentations des documents dans le but d'en récupérer des informations, à travers la construction d'index. L'informatique a permis le développement d'outils pour traiter l'information et établir la représentation des documents au moment de leur indexation, ainsi que pour rechercher l'information. On peut aujourd'hui dire que la recherche d'information est un champ transdisciplinaire, qui peut être étudié par plusieurs disciplines, approche qui devrait permettre de trouver des solutions pour améliorer son efficacité \*.

[\(\\*\) wikipedia](#)

## → Les étapes du processus de recherche d'information

<p><b>Étape 1</b>  <b>Je cerne le sujet</b></p>	<ul style="list-style-type: none"> <li>● Je comprends la nature, les objectifs et l'étendue du travail à exécuter</li> <li>● Je cerne le besoin d'information</li> <li>● Je précise la question de recherche</li> <li>● Je dégage les idées importantes et les mots-clés</li> </ul>
<p><b>Étape 2</b>  <b>Je formule le besoin d'information</b></p>	<ul style="list-style-type: none"> <li>● Je formule le besoin d'information</li> <li>● Je réfléchis aux ressources dont j'aurai besoin</li> </ul>
<p><b>Étape 3</b>  <b>Je repère mes sources d'information</b></p>	<ul style="list-style-type: none"> <li>● Je liste les ressources d'information que je pourrais éventuellement utiliser</li> <li>● Je sélectionne les sources pertinentes dans lesquelles je vais faire ma recherche</li> </ul>
<p><b>Étape 4</b>  <b>J'identifie les outils de recherche</b></p>	<ul style="list-style-type: none"> <li>● J'identifie les outils de recherche d'information que je pourrais utiliser en fonction des sources d'information choisis.</li> <li>● Je sélectionne l'outil le plus adéquat pour faire ma recherche</li> </ul>
<p><b>Étape 5</b>  <b>J'interroge des sources d'information</b></p>	<ul style="list-style-type: none"> <li>● Je construis des équations de recherche à partir de mes mots-clés</li> <li>● J'amorce la recherche à l'aide de mes mots-clés soumis à l'outil de recherche choisi</li> <li>● Je juge de l'efficacité de mes équations de recherche et je les ajuste au besoin</li> </ul>
<p><b>Étape 6</b>  <b>Je repère et j'évalue la qualité des résultats à récupérer</b></p>	<ul style="list-style-type: none"> <li>● Je repère les documents et les ressources qui me semblent pertinents</li> <li>● J'identifie les éléments d'information nécessaires à mon travail</li> <li>● J'analyse les informations prélevées au regard de la question de recherche</li> <li>● J'évalue la qualité de l'information selon des critères posés</li> <li>● J'examine différents points de vue</li> <li>● Je synthétise les idées et les informations provenant de sources variées</li> <li>● Je confronte des opinions et je construis la mienne selon l'objectif du travail</li> <li>● Je note la référence et les caractéristiques des documents retenus</li> <li>● J'indique mes sources d'information et je respecte les règles de la vie privée</li> </ul>

## II-Sources d'information

- Pour trouver une information, on doit la chercher dans les endroits éventuels où elle peut résider, ce qui constitue les sources d'information.
- Les sources d'information sont généralement imprégnées par les types d'information qu'elles renferment (formelle-informelle, générale-spécialisée, privée-public, etc.)
- Internet comporte de plus en plus d'information
- Plusieurs sources d'information traditionnelles sont désormais présentes sur la toile
- Le Web constitue aujourd'hui la première et principale source d'information, visible et cachée.
- L'information doit être accessible de façon continue
- La sélection d'information pertinente est une véritable tâche à réussir

On retrouve plusieurs sources d'information sur le Web, ci-après une liste non exhaustive non catégorisée et non classée de ces sources :

- Les agences de presse
- Les journaux électroniques
- Les e-magazines
- Les radios en ligne
- Les bibliothèques
- Les sites institutionnelles
- Les journaux spécialisés
- Les revues électroniques
- Les encyclopédies, les dictionnaires
- Les forums
- Les blogs
- Les réseaux sociaux
- Les organisations privées et publics
- Les associations, les comités
- Les publications scientifiques
- Les fils d'actualités
- Les wikis
- Les rapports, les cours, les manuels
- Les podcasts
- etc

→ ces sources peuvent être classées suivant **la validité et la fiabilité** des informations qu'elles renferment :

- Les sources d'information **formelles** (publication scientifique, site institutionnel, etc)
- Les sources d'information **informelles** (blog, wiki, etc.)

→ ces sources peuvent aussi être classées suivant **la spécialisation et la complexité** de l'information:

- Les sources d'information **généralistes** (revue scientifique de médecine)
- Les sources d'information **spécialisées** (ressources pédagogique sur le Data Mining)

→ ces sources peuvent aussi être classées suivant **les outils utilisés pour assurer la transmission et la diffusion** de l'information

- Les sources d'information **traditionnelles** (agence de presse, journaux, etc.)
- Les sources d'information **2.0** ou sociales (réseaux sociaux, blog, wiki, etc.)

## →Web invisible

Certaines informations résidant dans des sources d'information spécifiques sur le web peuvent ne pas être directement exploitables car non accessibles, d'ailleurs cela peut constituer la majeure partie des ressources présentes en ligne. c'est ce qu'on appelle **le web invisible**.

Le « web invisible » ou « web profond » désigne la partie du web non accessible aux usagers du web car les outils de recherche classiques ne peuvent pas les indexer.

Exemple : les sources d'information protégées par une authentification, les sites mal référencées, les nouveaux sites, etc.

## →Qualité des sources d'information :

La qualité des sources d'information dépend de :

- sa **fiabilité** (disposition à diffuser des informations exactes, liées au processus de validation de l'information ç-à-d sa précision et son exactitude)
- son **autorité** (réputation dont elle jouit dans son domaine auprès des experts)
- son **objectivité** (représentation fidèle, exempte de préjugé)

# III- Outils de recherche d'information

→ Annuaires de recherche

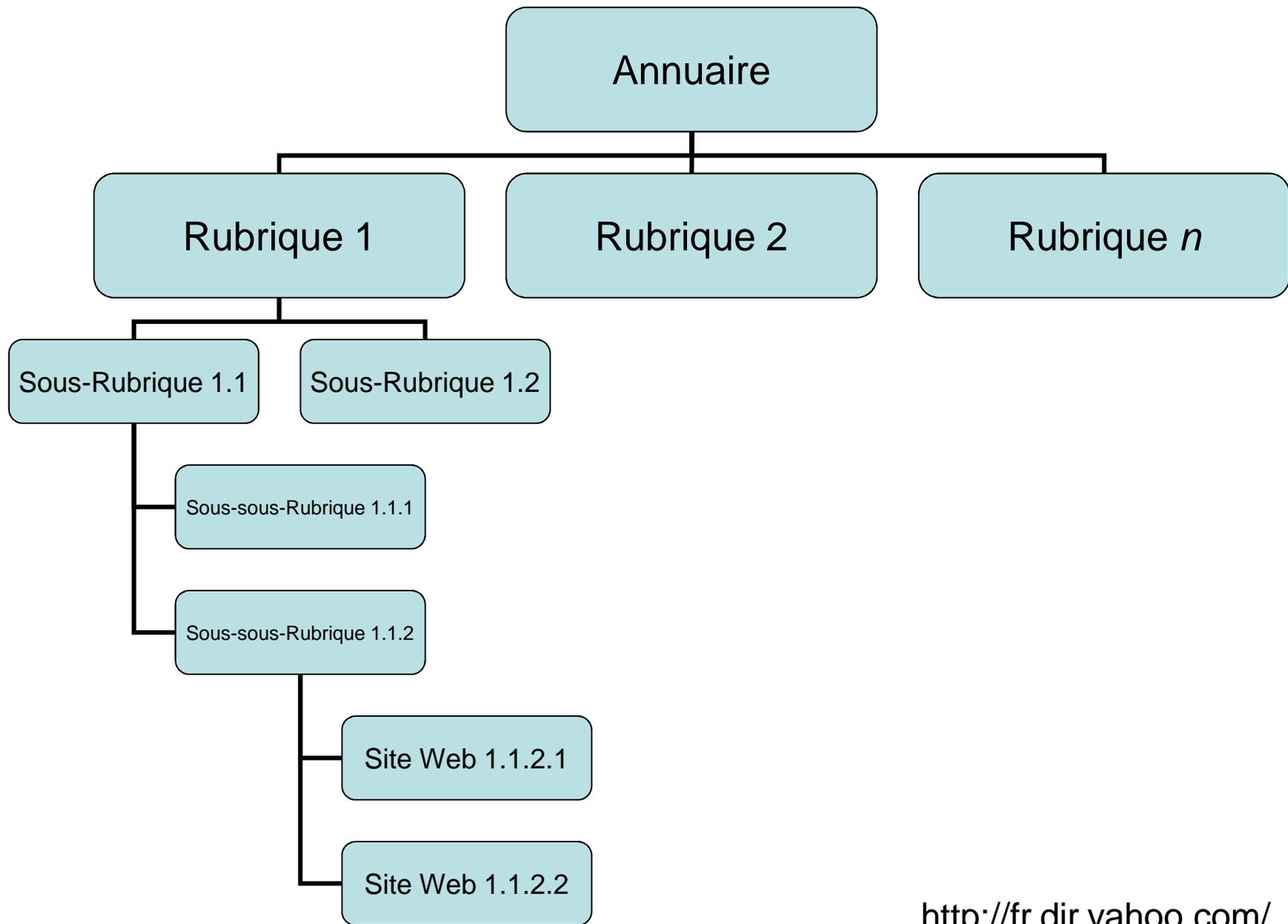
→ Moteurs de recherche

→ Meta-Moteurs

## → Les annuaires de recherche

### Définition :

- outil de recherche présentant une liste de sites web via une arborescence (classification, catégorisation, rubriques)
- classement en catégories et sous catégories
- sélection et ajout des sites webs par des professionnels



<http://fr.dir.yahoo.com/>

- Un annuaire (guide) est un outil de recherche permettant de répertorier et de recenser les sites Web.
  
- Un annuaire est organisé sémantiquement et hiérarchiquement, on trouve des rubriques, des sous-rubriques, des sous-sous-rubriques, etc.
  
- La dernière sous rubrique est une liste de sites Web (et non pas de pages Web).
  
- L'utilisation de l'annuaire de recherche se fait :
  - par l'arborescence et les thèmes (navigation)
  - par mots-clés : introduire le mot dans une boîte de recherche au niveau d'une rubrique ou sous-rubrique de l'annuaire. La recherche s'effectue sur les noms de rubriques de l'annuaire, les noms des sites et le bref descriptif ( non visible pour l'utilisateur ) associé à chaque site.

Annuaire Yahoo! - L'annuaire des sites Web de Yahoo! France - Mozilla Firefox

Fichier Édition Affichage Historique Marque-pages Outils ?

http://fr.dir.yahoo.com/

Les plus visités Débuter avec Firefox À la une KCC .. Kairouan-Cyber...

**YAHOO!** SEARCH

[Accueil Search](#) - [Yahoo!](#) - [Visite guidée](#) - [Aide](#)

## Annuaire des sites francophones de Yahoo! France

Go!

sur le Web  dans l'annuaire

Go!

[Recherche avancée](#)

### [Actualités et médias](#)

[Journaux](#), [Télévision](#), [Météo...](#)

### [Sports et loisirs](#)

[Foot](#), [Tourisme](#), [Auto/Moto](#), [Jeux...](#)

### [Commerce et économie](#)

[B2B](#), [Shopping](#), [Emploi](#), [Immobilier...](#)

### [Art et culture](#)

[Littérature](#), [Cinéma](#), [Musique](#), [BD...](#)

### [Informatique et Internet](#)

[Internet](#), [Logiciels](#), [Fonds d'écran...](#)

### [Divertissement](#)

[Tests/Quiz](#), [Loteries](#), [Humour](#), [Sorties...](#)

### [Santé](#)

[Diététique](#), [Médecine](#), [Thalasso...](#)

### [Classement géographique](#)

[Pays](#), [Europe](#), [France](#), [Paris...](#)

### [Enseignement et formation](#)

[Primaire](#), [Secondaire](#), [Supérieur...](#)

### [Références et annuaires](#)

[Dictionnaires](#), [Annuaire](#), [Cartes/Atlas...](#)

### [Institutions et politique](#)

[Ministères](#), [Droit](#), [Politique...](#)

### [Société](#)

[Enfants](#), [Gastronomie](#), [Rencontres...](#)

### [Sciences et technologies](#)

[Animaux](#), [Astronomie](#), [Physique...](#)

### [Sciences humaines](#)

[Archéologie](#), [Histoire](#), [Psychologie...](#)

Copyright © 2009 Yahoo! Tous droits réservés.

[Mise à jour](#) : [données personnelles](#) - [Conditions d'utilisation du service](#) - [Propriété Intellectuelle](#)

# YAHOO! DIRECTORY

Web |  Directory |  Category

## Education

[Email this page](#)

[Directory](#) > **Education**

SPONSOR RESULTS



[Earn Your Degree Online](#)

Ashford.Edu

Graduate sooner than you think!

CATEGORIES [\(What's This?\)](#)

### Top Categories

- [Adult and Continuing Education](#) (258)
- [Browse by Region](#) (169)
- [By Culture or Group](#) (234)
- [By Subject](#) (892)
- [Distance Learning](#) (840)
- [Higher Education](#) (17035)
- [K-12](#) (50192) **NEW!**

### Additional Categories

- [Academic Competitions](#) (66)
- [Bibliographies](#) (6)
- [Bilingual](#) (12)
- [Business to Business@](#)
- [Career and Vocational](#) (282)
- [Chats and Forums](#) (21)
- [Conferences](#) (22)
- [Correctional@](#)
- [Disabilities@](#)
- [Job and Employment Resources@](#)
- [Journals](#) (28)
- [Legislation](#) (6)
- [Literacy](#) (31)
- [News and Media](#) (178) **NEW!**
- [Organizations](#) (1911)
- [Policy](#) (40)
- [Programs](#) (294)
- [Reform](#) (94)

## - Construction

→ Indexation manuelle (humain)

## - Utilisation

→ Par navigation ou recherche par mots clefs

## - Avantages

→ Exploration simple

→ Sites sélectionnées

→ Précision et fiabilité de l'information

## - Inconvénients

→ Restrictions de taille

→ MAJ difficile

→ Sujets précis difficiles à trouver

# → Les moteurs de recherche

## Définition :

- Un moteur de recherche est un outil de recherche basé sur une **indexation automatique** des pages web et leur **référencement** dans une **base de données**
- Faire une recherche via un moteur de recherche revient à interroger la base de données en utilisant des **requêtes** et des **opérateurs booléens**
- Les pages non indexées (généralement non référencées) ne feront pas partie des résultats retournés (web invisible)
- Un moteur de recherche ne couvre pas la totalité du web.

- Un moteur de recherche est l'index d'un très grand nombre des pages du Web.
- Le fonctionnement d'un moteur de recherche est basé sur l'index.
- Un index fait correspondre pour chaque mot pertinent la page dans laquelle ce mot a été cité.
- La construction de l'index et son MAJ sont totalement automatiques.
- Google, Yahoo Search ou Bing sont des moteurs de recherche généralistes.
- Il existe aussi des moteurs de recherche spécialisés dans certains types de ressources (cartes, images, vidéos, publications scientifiques, blogs, etc.) ou dans certains domaines (immobilier, musique, films, médecine, etc.).
- Il existe aussi sur la plupart des sites web des moteurs de recherche dédiés aux pages des sites eux-même.

## Fonctionnement d'un moteur de recherche

- Des robots logiciels (appelés *crawlers* ou *spiders*) parcourent le Web, de lien en lien, de page en page ;
- Un outil parseur effectue la lecture des pages visités par le robot, détecte les principales parties et extrait les termes les plus pertinents ;
- L'indexeur construit l'index au fur et à mesure en ajoutant à chaque fois les termes et leurs références.

## Les limites de l'indexation automatique

- Certains documents numériques ne sont pas indexés car leurs formats ne sont pas pris en charge par le moteur.
- Les images et les vidéos sont indexés grâce aux métadonnées associées (données de description de ces ressources, les descriptions peuvent être extraites à partir des tags et données textuelles qui y sont associées)

## Recherche dans les moteurs de recherche

- L'utilisateur formule sa requête (ensemble de mots clefs)
- Le moteur de recherche :
  - consulte l'index,
  - recherche des entrées correspondantes  
au(x) mot(s) demandé(s)
  - Sélectionne les pages trouvées
  - Présente les adresses des pages suivant un ordre  
décroissant de pertinence

## Formulation de requêtes

- Principalement avec AND, OR
- Mais aussi avec NOT, NEAR, SAUF

## -Avantages

- Taille d'index est proportionnelle à la taille du Web
- MAJ fréquente
- Grand nombre de résultats retournés

## -Inconvénients

- Formulation de la requête
- Nombre de résultats souvent très énorme
- Recherche automatique, certaines pages trouvées n'ont pas de relation sémantique avec la requête

Google - Mozilla Firefox

Fichier Édition Affichage Historique Marque-pages Outils ?

http://www.google.com/

Les plus visités Débuter avec Firefox À la une KCC .. Kairouan-Cyber...

Web [Images](#) [Maps](#) [Actualités](#) [Vidéo](#) [Gmail](#) [plus](#) ▼

[iGoogle](#) [Connexion](#)



Rechercher sur le Web  Rechercher les pages en français

[Recherche avancée](#)  
[Préférences](#)  
[Outils linguistiques](#)

[Programmes de publicité](#) - [Solutions d'entreprise](#) - [À propos de Google](#) - [Google.com in English](#)

©2009 - [Confidentialité](#)



## Recherche avancée

[Conseils de recherche](#) | [À propos de Google](#)

<b>Pages contenant</b>	tous les mots suivants	<input type="text"/>	10 résultats	<input type="button" value="Recherche Google"/>
	cette <b>expression exacte</b>	<input type="text"/>		
	<b>au moins un</b> des mots suivants	<input type="text"/>		
	<b>aucun</b> des mots suivants	<input type="text"/>		
<b>Langue</b>	Résultats pour les pages écrites en	toutes les langues		
<b>Région</b>	Rechercher les pages disponibles en :	tous les pays/territoires		
<b>Format de fichier</b>	Seulement afficher les résultats avec des fichiers au format	tous les formats		
<b>Date</b>	Rechercher les pages Web découvertes durant	date indifférente		
<b>Emplacement</b>	Pages dans lesquelles le ou les termes figurent	n'importe où dans la page		
<b>Domaines</b>	Seulement afficher les pages du site ou du domaine	<input type="text"/>		
		par exemple google.com, .org, .fr, etc. <a href="#">Plus de détails...</a>		
<b>Droits d'utilisation</b>	Afficher les résultats	sans tenir compte des licences		
<b>SafeSearch</b>	<input checked="" type="radio"/> Aucun filtrage <input type="radio"/> Filtre <a href="#">SafeSearch</a>			

### Recherche de pages spécifiques

<b>Similaires</b>	Rechercher les pages similaires à cette page	<input type="text"/>	<input type="button" value="Rechercher"/>
		par exemple : www.google.com/help.html	
<b>Liens</b>	Rechercher les pages liées à cette page	<input type="text"/>	<input type="button" value="Rechercher"/>

### Recherches sur un sujet précis

Nouveau ! [Google Recherche de code](#) - Recherche de code source public

## → Les méta-moteurs

### **Définition :**

- Un métamoteur de recherche est un outil de recherche basé sur l'utilisation d'une série de moteurs de recherche et/ou annuaires de recherche ;
- Il interroge simultanément plusieurs moteurs et/ou annuaires de recherche ;
- Il récupère les résultats et les classe suivant l'indice de pertinence ou autres.

### **Pourquoi avoir recours à un métamoteur de recherche ?**

- accéder à un plus grand nombre de pages référencées sur Internet ;
- Bénéficier d'un plus grand nombre de fonctionnalités ;
- avoir un plus grand nombre de résultats ;
- améliorer la recherche ;
- avoir plus de chances à trouver des documents plus pertinents.

# metacrawler®

▶ SEARCH THE SEARCH ENGINES!®

Including Google, Yahoo!, MSN Search and Ask. [Learn More](#)

Web | [Images](#) | [Video](#) | [News](#) | [Yellow Pages](#) | [White Pages](#)

[Advanced Search](#) | [Preferences](#)

### Popular Searches

<a href="#">spring clean</a>	<a href="#">camping gear</a>
<a href="#">iphone</a>	<a href="#">ps3 cheats</a>
<a href="#">mp3 downloads</a>	<a href="#">sports news</a>

[Make MetaCrawler Your Homepage](#)

[Add Metacrawler Search to Your Site](#)

[About MetaCrawler](#) | [Privacy Policy](#) | [Terms of Use](#) | [Contact Us](#) | [Tools & Tips](#)

[About InfoSpace](#) | [Doopile Search](#) | [WebCrawler Search](#) | [Nation Search](#) | [Careers](#) | [Press](#) | [Investor Relations](#)

© 2009 InfoSpace, Inc. All Rights Reserved

## IV- Formulation de requête

- Une requête de recherche bien formulée garantit un résultat de recherche de qualité (précision, pertinence, temps de réponse..)
- Formuler une requête de recherche nécessite la maîtrise d'un langage d'interrogation spécifique
- Le langage d'interrogation utilisé dans la plupart des outils de recherche d'information est basée sur l'adjonction d'une série de mots clefs représentant le besoin d'information de l'utilisateur

## Stratégie de recherche

- Bien se fixer le besoin d'information
- Traduire ce besoin en une ou plusieurs questions
- Formuler ces questions en langage courant d'interrogation
- Extraire les principaux concepts de ces questions
- Utiliser des mots clefs et leurs synonymes
- Relier ces mots clefs par des opérateurs booléens pour que la recherche soit bien ciblée

# Les opérateurs de recherche

→ OU / OR



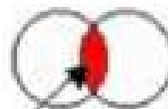
"OU" EST L'OPÉRATEUR DE L'UNION - SPÉCIFIQUE

→ SAUF / NOT



LE « SAUF » EST L'OPÉRATEUR DE L'EXCLUSION - SPÉCIFIQUE

→ Et / And



LE « ET » EST L'OPÉRATEUR DE L'INTERSECTION - SPÉCIFIQUE

# Troncature

→ Pour remplacer un nombre de caractères après la racine (début) du mot\*, on utilise le caractère étoile « \* » .

→ Le caractère « ? » est utilisé pour ne remplacer qu'un seul caractère dans le mot à soumettre au moteur de recherche.

→ (\*) La troncature peut être utilisée :

- à droite d'un mot ;
- à gauche d'un mot ;
- à l'intérieur d'un mot.

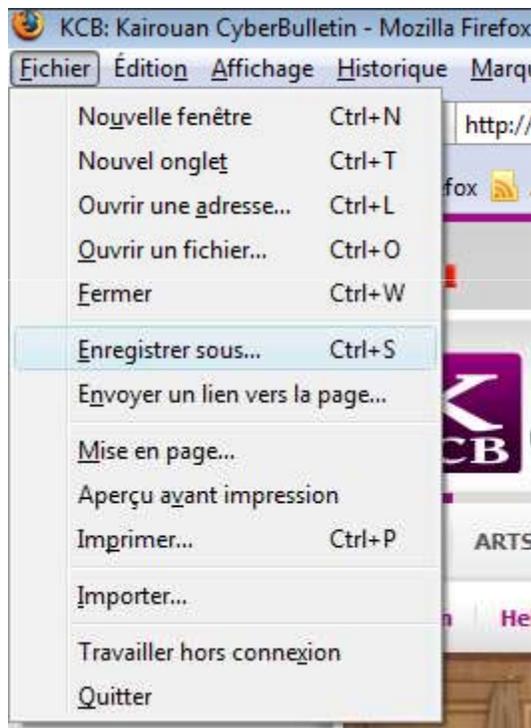
## Recherche d'expressions exactes

→ Pour rechercher les documents contenant exactement une expression, on place les expressions à chercher entre guillemets "...".

→ A utiliser pour une requête dans laquelle les termes ne sont pas recherchés séparément (mot par mot), mais comme une expression.

# V- Récupération des informations

→ Récupérer une page HTML

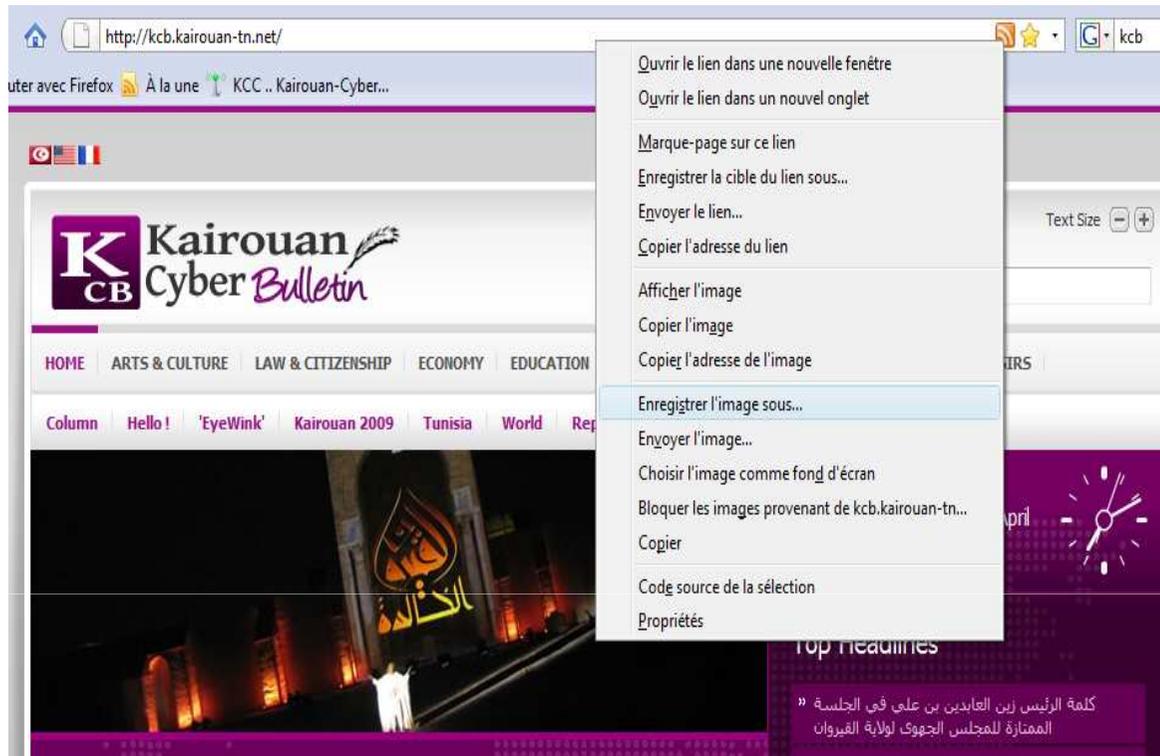


-Sélectionner le menu « Fichier » et choisir l'option « Enregistrer sous »

-Choisir ensuite l'emplacement dans lequel le contenu de la page HTML sera sauvegardé

-Des dossiers images seront créés dans l'emplacement d'enregistrement afin de sauvegarder les images de la page

## → Enregistrer une image



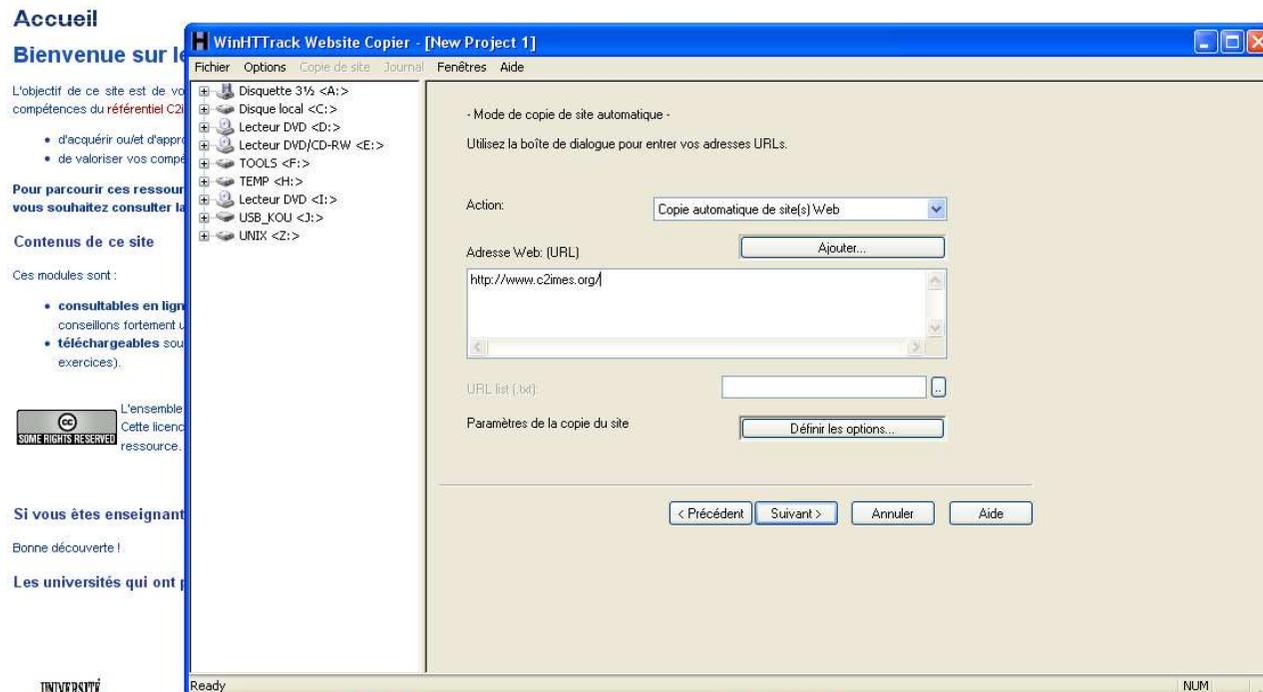
- Cliquer sur l'image à enregistrer en utilisant le bouton droit de la souris et choisir l'option « Enregistrer l'image sous »
- Choisir ensuite l'emplacement dans lequel l'image sera sauvegardé

## → Aspirer un site

- Télécharger l'ensemble du site Web en reproduisant la même arborescence sur disque dur pour permettre une navigation offline
- Plusieurs outils d'aspiration Web (gratuits et payants) peuvent être téléchargés depuis Internet.



Exple :  
WinHTTrack



Fin ...



جامعة تونس الافتراضية  
UNIVERSITE VIRTUELLE DE TUNIS

CERTIFICAT INFORMATIQUE & INTERNET



## **D4-2 : Evaluer les résultats d'une recherche**

Préparé par : Mohamed Koutheaïr KHRIBI

# Plan

I-Introduction

II-Critères d'évaluation

III-Evaluation d'une ressource web

# I-Introduction

→ Suite à une opération de recherche d'information, l'on peut bien se poser maintes questions sur la nature des résultats trouvés : sont-ils satisfaisants ? sont-ils corrects ? sont-ils précis ? sont-ils exacts ? sont-ils fiables ? sont-ils pertinents ? sont-ils intéressants ? Finalement est ce que ça représente bien ce que l'on cherchait vraiment ?

→ Toutes ces questions sont légitimes voire-même nécessaires et obligatoires et ce, avant de passer à l'étape de l'utilisation des résultats de recherche d'information, il s'agit donc de les évaluer !

→ Pour évaluer des résultats de recherche, il nous faut des critères d'évaluation, qu'est ce qu'on va évaluer, suivant quels critères et comment l'évaluation aura lieu ?

## II- Critères d'évaluation

-Principalement, l'on peut se baser sur les deux critères suivants pour assurer l'évaluation d'un résultat de recherche d'information :

→La pertinence

→La fiabilité

- La **pertinence** d'une ressource est l'adéquation de son contenu informationnel avec le besoin d'information.

-La **fiabilité** de l'information est la confiance qu'on peut accorder à sa véridicité.

# II- Evaluation d'une ressource Web

Pour évaluer une ressource Web, les règles suivies reposent sur les questions suivantes :

→ Qui? - QUOI? - OU? - QUAND? - COMMENT? - POURQUOI?

## **QUI ?**

L'auteur d'une ressource est son premier responsable. L'identifier, déterminer ses compétences, ses activités, son affiliation, évaluer sa notoriété, etc. s'avère très important.

## QUOI ?

→Quelle est la nature de la ressource (évaluer la source d'information) ?  
(formel/informel, traditionnel/social, public/privé, etc.)

→Y a-t-il des liens vers des sites fiables ? La ressource est-elle pointée par des sites fiables?

→L'information fournie existe-t-elle sur d'autres ressources? Est-elle réservée uniquement à la ressource trouvée ?

→A quel public s'adresse-t-on (spécialiste, général, connaisseur, non-initié, ...)?Le traitement du sujet est-il exhaustif?

## Où?

→Par quel serveur le site est-il hébergé? Sous quelle forme l'URL du site parait-elle (domaine, sous-domaine, nom de domaine, etc.) et ce, pour avoir un idée sur la nature du site, son origine, sa nationalité, son statut, sa catégorie, etc.

→Le public-cible peut-il accéder facilement au site?

## **QUAND ?**

→ Il s'agit d'évaluer la fraîcheur de l'information en consultant la date de création de la ressource, sa date de mise à jour, la fréquence de mise à jour, etc.

→ Quelle est la date de création du site? Quelle est celle de sa dernière mise à jour? Cette date est-elle affichée dans la page? Les liens vers d'autres sites sont-ils à jour? Existe-t-il beaucoup de liens morts?

## **COMMENT?**

→ Comment se présente la ressource? Comment l'information est-elle rédigée? Comment la ressource est-elle structurée? Y a-t-il des sources, sont-elles bien indiquées?

→ Comment accède-t-on à l'information? L'accès est-il public ou privé? Est-il payant ou gratuit? La navigation est-elle aisée? Le chargement du site est-il rapide?

## POURQUOI ?

→Évaluer l'objectif de la ressource : Le type de la ressource renvoie souvent aux objectifs correspondants.

→Il faut bien ainsi révéler et cerner les objectifs visés par la ressource

## → **Evaluation de la fiabilité d'une ressource**

- La validité d'une ressource peut nous aider à déterminer sa fiabilité.
- Une ressource peut ne pas avoir de validité (inexistante) –c'est le cas des blogs, sites d'opinion, sites personnels, commerciaux etc.);
- Une ressource peut être validée suivant la démarche traditionnelle ç-à-d s'il s'agit de sites spécialisés, institutionnels, revue scientifique, publication académique, etc.
- Une ressource peut être validée (éventuellement) suivant la démarche sociale ou collaborative et ce, en étant éditée, partagée et approuvée par un ensemble d'auteurs et en respectant un nombre de règles. Ce qui est assez original. (c'est le cas des wiki, en l'occurrence wikipedia)

Fin ...



جامعة تونس الافتراضية  
UNIVERSITE VIRTUELLE DE TUNIS

CERTIFICAT INFORMATIQUE & INTERNET

## **D4-3 : Récupérer et référencer une ressource numérique en ligne**

Préparé par : Mohamed Koutheaïr KHRIBI

# Plan

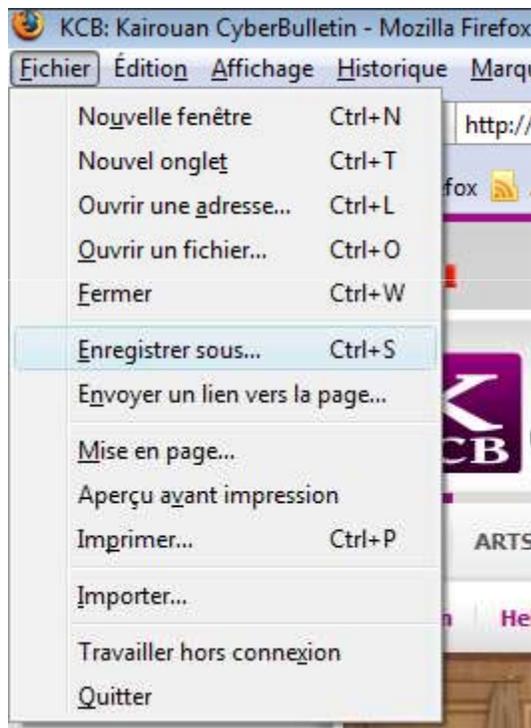
I-Récupération d'information

II-Disponibilité de ressource

III-Référence de ressource

# I- Récupération des informations

→ Récupérer une page HTML

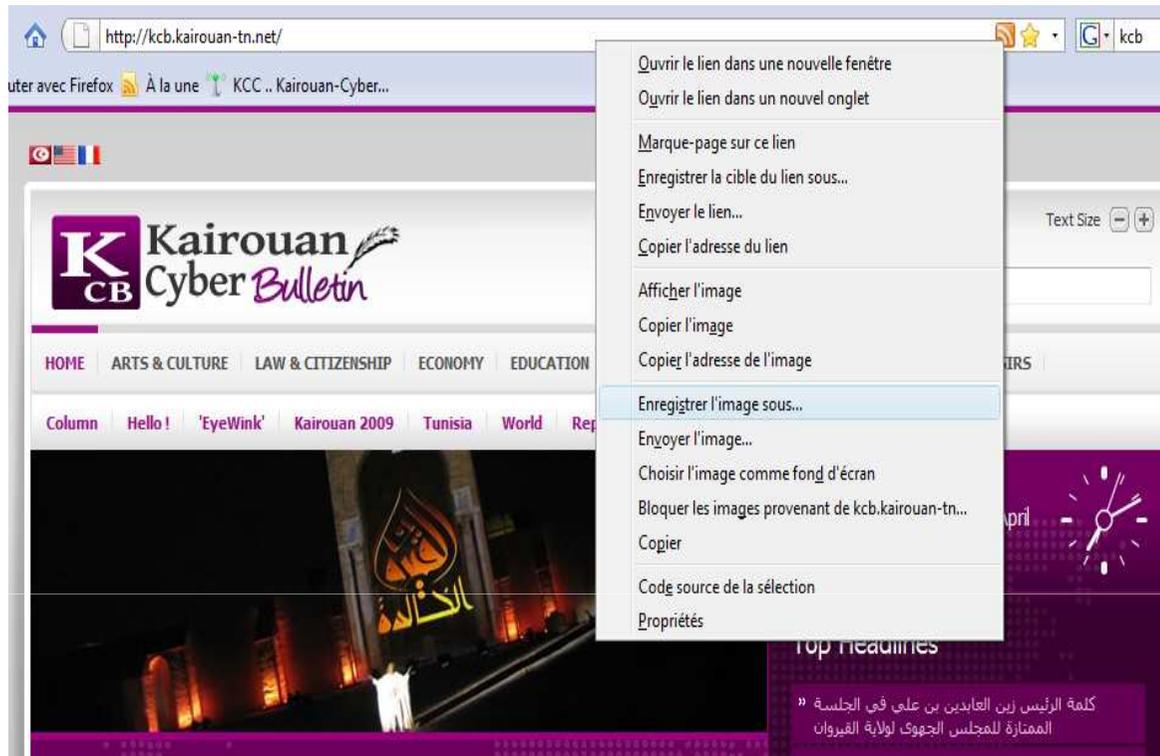


-Sélectionner le menu « Fichier » et choisir l'option « Enregistrer sous »

-Choisir ensuite l'emplacement dans lequel le contenu de la page HTML sera sauvegardé

-Des dossiers images seront créés dans l'emplacement d'enregistrement afin de sauvegarder les images de la page

## → Enregistrer une image



- Cliquer sur l'image à enregistrer en utilisant le bouton droit de la souris et choisir l'option « Enregistrer l'image sous »
- Choisir ensuite l'emplacement dans lequel l'image sera sauvegardé

→ Récupérer une partie de la page, copier un texte par exemple :

Newsletter

Home Tek Net Phone Médias Ness Kult

Suivez nous partout recherche...

Tek | Tek News | Mercredi, 30 Novembre 2011 07:05

## Tunisie : Le e-commerce au salon, en attendant sa diffusion

[0 Comments](#)

Le salon e-Commerce permettra de faire le point sur les nouvelles tendances d'un domaine prometteur, même s'il n'a pas encore laissé exprimer toute l'étendue de son potentiel en Tunisie. 18 conférences thématiques sont prévues, avec à la clé des professionnels internationaux du e-commerce et du e-marketing.



Le salon e-Commerce Tunis ouvrira ses portes le 1 et le 2 Décembre 2011 à l'hôtel Le Palace à Gammarth, dans la banlieue nord de Tunis. Pour sa 1ère édition, cet événement organisé par Xpert-Net, a réservé un espace d'exposition de 1 000 m<sup>2</sup> en regroupant les intervenants tunisiens et étrangers du secteur.

Des spécialistes, et des experts témoigneront de leurs expériences au cours de leur intervention. Des professionnels du e-commerce et du e-marketing, interviendront ainsi pour encourager, aider les hommes d'affaires et les citoyens à sauter le pas. Des acteurs internationaux du secteur participeront à des forums dédiés au e-banking, au e-tourisme et au m-mobile.

Le salon permettra de faire le point sur les nouvelles tendances d'un domaine prometteur, même s'il n'a pas encore laissé exprimer toute l'étendue de son potentiel en Tunisie. Le marketing en ligne, le e-commerce représentent en effet un gisement potentiel d'emplois à ne pas négliger dans le contexte actuel.

Ainsi, des conférences et des ateliers seront centrés sur les nouvelles technologies, pour familiariser le grand public avec les enjeux du commerce électronique. Ce forum offrira donc l'opportunité de nouer des contacts

NEWS

- Tunisie Telecom : Payez la facture avant sa réception
- Tunisie : Laboratoire des Interfaces et Matériaux à Monastir
- Twitter s'arabise
- LG lance un nouveau climatiseur respectueux de l'environnement
- Tunisie : L'Open Source arabe en question

les bonnes affaires sur **Lesannonces.tn**

- Annonces immobilier Tunisie
- Achat appartement Tunisie
- Annonces auto Tunisie

www.lesannonces.tn

ESPACE VIDEOS

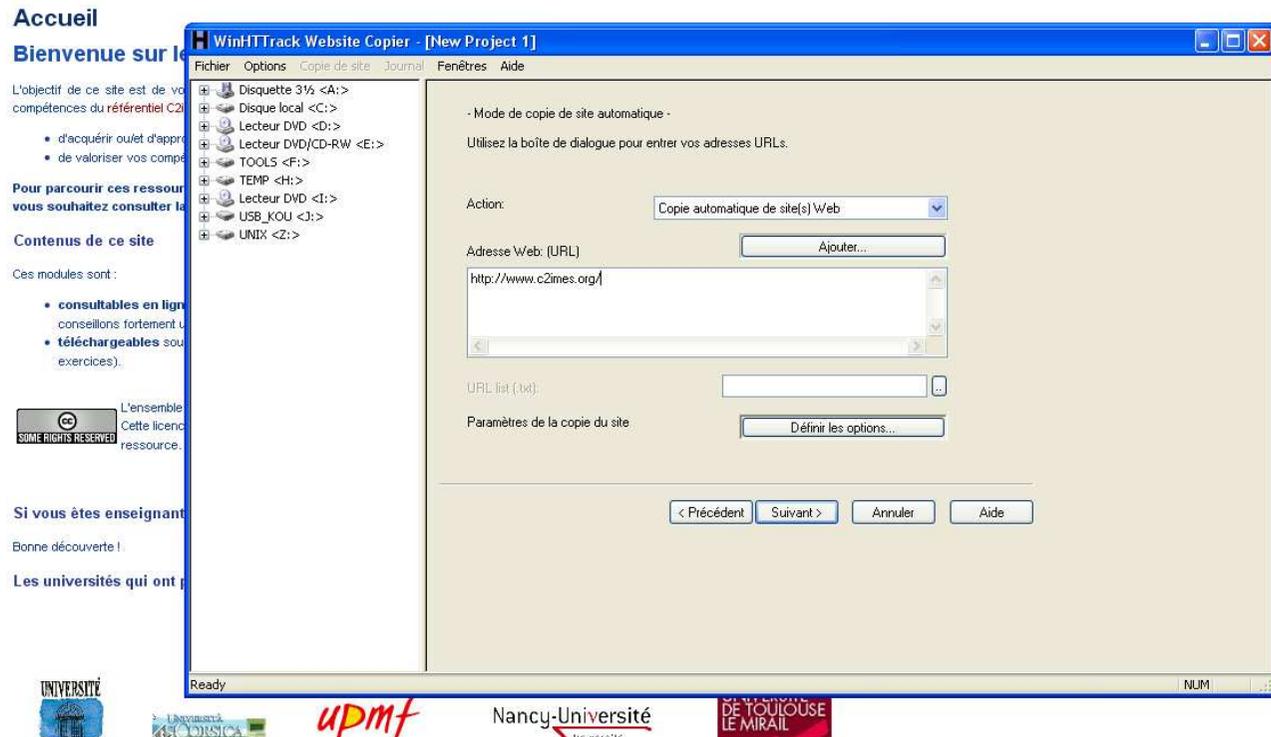
Sélectionner la partie de la page à récupérer, par exemple un bout du texte de l'article, le copier et le coller dans un autre document.

## → Aspirer un site

- Télécharger l'ensemble du site Web en reproduisant la même arborescence sur disque dur pour permettre une navigation offline
- Plusieurs outils d'aspiration Web (gratuits et payants) peuvent être téléchargés depuis Internet.



Exple :  
WinHTTrack  
<http://www.httrack.com>



## II-Disponibilité de ressource

Après avoir trouvé la bonne ressource web et après avoir effectué sa récupération pour éventuellement la réutiliser dans un contexte particulier (rapport, présentation, étude, etc.), il ne faut pas oublier de citer cette ressource suivant les règles de l'art.

Néanmoins, il faut faire attention que les liens cités ne soient obsolètes car le web contient de plus en plus de ressources devenues éphémères ou instables étant donné que leur contenu change fréquemment au bout du temps ou bien l'inverse ç-à-d devient non utilisé et non entretenu, on parle dans ce cas de lien mort ou aveugle.

Par exemple, dans le cadre d'un site web dynamique institutionnel ou un site de journal électronique ou agence de presse ou wiki, on peut trouver un article dont le contenu est modifié entre le moment où on l'a consulté pour y faire référence dans un rapport, et le moment où le lecteur du rapport le consulte.

→ Ainsi, lorsqu'on fait référence à ces ressources, il faut être conscient de ce genre de difficultés (changement de contenu, lien mort, changement d'adresse, etc.) et prendre les mesures nécessaires.

**→ Faire attention au moment de la :**

--| Copie de l'adresse URL de la ressources, dans le cas où on va copier l'adresse d'un article particulier d'un journal électronique, il faut copier toute l'adresse (longue) de l'article et non pas seulement l'adresse du journal, par exemple copier : <http://www.tekiano.com/tek/4-tek-news/4530-tunisie-le-e-commerce-au-salon-en-attendant-sa-diffusion.html> et non pas <http://www.tekiano.com>

--| la date de consultation : toujours indiquer la date de consultation avec précision afin de mettre l'article dans son cadre temporel d'apparition ou de publication

--| la consultation : ne pas oublier de faire une sauvegarde de la ressource consultée pour en garder toujours une copie du contenu original, ce qui permet de retrouver à coup sûr la ressource pour vérification ou éventuellement preuve

## II-Référence de ressource

Pour citer une ressource web, utilisée par exemple dans un rapport dans le cadre des références bibliographiques, suivant les règles de l'art, il faut ajouter des renseignements spécifiques à cette ressource. Pour ce faire, il existe des normes contenant des règles spécifiques à respecter.

Pour la citation des documents électroniques en particulier, il existe une norme définissant la composition et la typographie d'une référence bibliographique (référence ISO 690-2 ou Z 44-005-2).

Les règles sont les mêmes que pour les documents traditionnels, mais il faut ajouter :

- le type de support (en ligne, CD-ROM, enr. sonore, etc.) ;
- l'adresse de la ressource (URL) ;
- la date de consultation de la ressource.

## → Cas d'un site web (en général)

Modèle :

**Auteur. Date. Titre [Support]. Adresse de la ressource. (consulté le ...)**

- Auteur : l'organisme ou la personne physique dans le cas d'une page personnelle.
- Date : celle de la dernière mise à jour si elle est connue.
- Titre : celui de la page d'accueil du site.
- Support : il s'écrit entre crochets ; on trouve : [En ligne] [CD-ROM] [Enr. sonore], etc.
- Adresse de la ressource : c'est l'URL.
- Date de consultation par l'utilisateur : elle s'écrit (consulté le jour mois année).

### Exemple :

E-magazine Tekiano. (Mercredi, 30 Novembre 2011). Accueil du site [En ligne].

<http://www.tekiano.com> (consulté le 30/11/2011).

## → Cas d'une ressource Internet

(précise cas d'un article ou d'une page web en particulier):

**Auteur. Date. «Titre de la ressource». In : Ressource plus large, [Support]  
Adresse de la ressource (consulté le ...)**

- Ressource plus large : s'il y a lieu, ajouter la ressource plus large à laquelle le document cité est rattaché; procéder de la façon suivante: In Auteur/titre ou titre du site ou du document qui contient la ressource.

### **Exemple :**

E-magazine Tekiano. (Mercredi, 30 Novembre 2011 07:05) Tunisie :

Le e-commerce au salon, en attendant sa diffusion. In : <http://www.tekiano.com>

[En ligne]. [http://www. http://tekiano.com/tek/4-tek-news/4530-tunisie-le-e-commerce-au-salon-en-attendant-sa-diffusion.html](http://www.tekiano.com/tek/4-tek-news/4530-tunisie-le-e-commerce-au-salon-en-attendant-sa-diffusion.html) (consulté le 30/11/2011).

Fin ...



CERTIFICAT INFORMATIQUE & INTERNET

جامعة تونس الافتراضية  
UNIVERSITE VIRTUELLE DE TUNIS



## **D4-4 : Organiser une veille informationnelle**

Préparé par : Mohamed Koutheaïr KHRIBI

# Plan

I-Veille informationnelle

II-Outils Web utilisés

# I-Veille informationnelle

## →Définition :

La veille informationnelle consiste à surveiller couramment la parution d'information pertinente afin de se tenir toujours au courant de l'actualité concernant un domaine ou d'un sujet particulier. La veille informationnelle est assurée en faisant des actions de recherche, de traitement et de diffusion.

La veille informationnelle peut se faire dans un cadre personnel dans le but d'être à la page en ce qui concerne un sujet particulier (domaine de commerce, recherche universitaire, mode, etc.) ou dans un cadre professionnel afin d'exploiter les informations recueillies et en tirer profit, ce qui représente un avantage concurrentiel.

## → Outils:

Pour faciliter le travail de veille informationnelle, on peut utiliser un nombre d'outils informatiques issus du domaine des nouvelles technologies, en l'occurrence les outils Web 2.0 .

Par exemple, on peut utiliser :

- Un outil d'agrégation des flux d'actualités diffusés par différents sites web, ce qui nous ménage la visite répétitive de ces sites un par un pour voir les nouveautés puisque l'outil permet d'afficher automatiquement les dernières actualités dans un même espace
- Un outil de Newsletter, ce qui permet de s'abonner à des lettres d'information afin que les nouveautés et les actualités d'un site web nous parviennent par messagerie électronique de façon hebdomadaire ou mensuelle.
- Un outil de billets à travers les blogs. En effet, on peut suivre un ou plusieurs veilleurs sur un système de blogage, ceci nous permettra de suivre des veilleurs spécialisés dans différents domaines et en choisir celui qui nous intéresse pour être au courant de toutes les actualités dans un domaine particulier. Ainsi, un filtrage et une sélection au préalable est déjà faite sans aucun effort.
- Un outil de forum de discussion avec l'utilisation d'un système de notification par messagerie électronique pour nous ménager l'accès et la consultation quotidienne du forum. En effet, quand un nouveau message est inséré, un email simultanément contenant une copie du message est envoyé à l'ensemble des abonnés, on est alerté lorsqu'une modification a été apportée.

# II-Outils Web utilisés

La plupart des outils utilisés pour assurer la veille informationnelle sont des outils Web 2.0.

## **Qu'est ce que le Web 2.0**

Dans sa forme et usage d'origine, le Web, nommé le « Web 1.0 », comprenait essentiellement des pages web statiques qui étaient rarement mises à jour, voire jamais. Par la suite, les sites web ont pu intégrer des informations recueillies de façon dynamique à travers l'utilisation des bases de données et notamment grâce à l'avènement des systèmes de gestion de contenu CMS . Ces derniers ont permis une démocratisation de la création du contenu web et sont à l'origine du début de changement du rôle de l'internaute qui était jusque là plutôt passif par rapport au Web considéré principalement comme un outil de diffusion et de visualisation de données.

Aujourd'hui, l'heure est maintenant à l'interactivité avec l'internaute dont la création du contenu web n'est plus un secret pour lui, on parle même d'un phénomène de prise de pouvoir de l'internaute.

Cette nouvelle approche du web est qualifiée d'approche « Web 2.0 » où l'utilisation du Web s'oriente donc de plus en plus vers l'interaction entre les utilisateurs qui ont la possibilité, à travers d'une panoplie d'outils, de créer facilement du contenu, d'en discuter, d'y ajouter des commentaires, de communiquer et de collaborer avec d'autres utilisateurs, etc.

Le terme « Web 2.0 » a été inventé par Dale Dougherty de la société O'Reilly Media

lors d'un « brainstorming » avec Craig Cline de MediaLive pour développer des idées

pour une conférence conjointe. Il a suggéré que le Web était dans une période de

renaissance, avec un changement de règles et une évolution des « business model ».

Les approches « Web 2.0 » les plus populaires sont sans aucun doute les phénomènes « blog », « wiki » et « réseaux sociaux ». Les blogs sont basés sur un principe de rédaction individuelle annotée par les autres alors que les wikis sont basés sur un principe de rédaction collective. Les réseaux sociaux représentent une sorte de communauté d'individus ou d'organisations reliés de manière directe ou indirecte entre eux, en fonction de centres d'intérêts, de points de vue ou encore de besoins communs.

## Définition

« L'expression « Web 2.0 » désigne certaines technologies et des usages du World Wide Web qui ont suivi la forme initiale du web, en particulier les interfaces permettant aux internautes ayant peu de connaissances techniques de s'approprier les nouvelles fonctionnalités du web. Ainsi, les internautes contribuent à l'échange d'informations et peuvent interagir (partager, échanger, etc.) de façon simple, à la fois avec le contenu et la structure des pages, mais aussi entre eux, créant ainsi notamment le Web social. L'internaute devient, grâce aux outils mis à sa disposition, une personne active sur la toile. »  
*(wikipedia)*

# Technologies

L'infrastructure du Web 2.0 est complexe et changeante, mais elle inclut les logiciels de serveur, la syndication de contenu, les protocoles de messagerie, des standards de navigation, et des applications clientes diverses. Ces approches complémentaires fournissent au Web 2.0 les capacités de stockage, de création et de diffusion qui vont au-delà de ce qui était précédemment attendu des sites Web.

Un site pourrait être appelé comme utilisant une approche Web 2.0 s'il adopte un certain nombre des techniques, telles que :

- l'utilisation de CSS, d'un balisage XHTML sémantiquement valide et des microformats ;
- les techniques d'applications riches telles qu'AJAX ;
- la syndication et l'agrégation de contenu RSS/Atom ;
- la catégorisation par étiquetage ;
- l'utilisation appropriée des URL ;
- l'utilisation des services Web XML.

## Flux RSS (agrégateur de flux)

RSS (Really Simple Syndication) est un format de description de données qui facilite la diffusion de fils d'informations, de façon automatisée, à l'intention des internautes, et qui favorise la syndication de contenus en permettant à d'autres sites de republier facilement tout ou partie de ces données.

Grâce à ce format, plus besoin de consulter directement ses sites préférés pour être informé de leurs nouveautés : il suffit à l'internaute concerné d'insérer les fils d'informations RSS dans un logiciel compatible pour avoir toutes les sources d'informations agrégées sur un seul écran !

Pour accéder aux flux RSS, Il existe plusieurs moyens :

- utiliser un logiciel de messagerie et profiter des informations agrégées en même temps que la consultation des e-mails,
- utiliser un navigateur Web et afficher les différents flux d'informations sous forme d'une page d'accueil,
- utiliser un logiciel spécialisé dans l'agrégation de fils d'informations RSS, également connu sous le nom de lecteur RSS ou encore d'agrégateur d'actualités,
- ou encore profiter de ces fils d'informations sur un Smartphone, un Palm ou un Pocket PC connecté en Bluetooth à un mobile en GRPS.

Quelle que soit la formule retenue, pour utiliser les flux RSS, il faudra inscrire dans le logiciel utilisé l'adresse URL de chaque flux d'informations.

En effet, la plupart des sites diffusant des fils d'informations affichent une petite icône sur leurs pages de type RSS ou XML, icône sur laquelle il suffit de cliquer pour que le fil d'informations soit automatiquement reconnu par votre logiciel et ajouté à votre agrégation de contenus.



## Principaux agrégateurs RSS

- RSS Owl est un lecteur RSS multi-plateformes (Windows, Linux, Mac et Solaris) en Java. Il permet notamment l'import et l'export au format OPML :
- RSS Reader est un lecteur pour Windows, nécessitant l'installation du Framework .NET
- FeedReader est un lecteur RSS supportant les formats RSS et Atom :
- Le client de messagerie Mozilla Thunderbird permet nativement d'exploiter les fils RSS et de les afficher comme du courrier
- Le navigateur Mozilla Firefox permet de s'abonner à un fil RSS et d'afficher les actualités via ses menus déroulants
- Il existe également des extensions pour Firefox (lecteurs RSS améliorés)

# Blog

Un «blog» (en français blogue, journal web ou carnet web) est un site internet personnel permettant à un internaute de publier facilement des informations en ligne sous forme de journal. Le terme «blog» est la contraction de «weblog» (au même titre, «blogue» est la contraction de «weblogue»).

Le blog est donc une page web, s'appuyant généralement sur un outil dynamique permettant de gérer la mise en page du contenu, dans laquelle son auteur publie des actualités, des liens vers des sites web ou de courts textes de sa création.

Le dernier article publié est toujours en première page du blog.

A noter que les blogs sont en général gérés comme des CMS.

Un blogueur a aujourd'hui loisir de mélanger textes, hypertexte et éléments multimédias (image, son, vidéo, applet) dans ses billets ; il peut aussi répondre aux questions et commentaires des lecteurs car chaque visiteur d'un blog peut ou bien laisser des commentaires sur le blog lui-même, ou bien contacter le blogueur par courrier électronique.

## Tableau de bord

Langue : Français ▼

**Nouveau !**

Nous avons activé la détection automatique de spam pour les commentaires. Pensez à consulter régulièrement le dossier des messages indésirables. [En savoir plus sur la fonction de détection du spam de Blogger](#) ou [signaler des problèmes](#).

koutheair

Gérer les blogs (1 au total)

[Créer un blog](#) - [Aide](#)[Afficher le profil](#)[Modifier mon profil](#)[Modifier la photo](#)[Modifier les notifications](#)  **koutheair**1 message, dernière publication le 17 janv. 2009 - [Afficher le blog](#) [NOUVEAU MESSAGE](#) - [Modifier les messages](#) - [Commentaires](#) - [Paramètres](#) - [Présentation](#) - [Monétisation](#) - [Statistiques](#)

## Liste de lecture

Les blogs dont je suis membre

Blogger Buzz

Blogs intéressants

Blogs

Éléments

Toutes les mises à jour du blog

Bienvenue dans la Liste de lecture. Voici toutes les mises à jour pour 

# Wiki

Un wiki est un système de gestion de contenu de site Web qui rend les pages Web

librement et également modifiables par tous les visiteurs autorisés.

Les wikis sont utilisés pour faciliter l'écriture collaborative de documents avec un minimum de contrainte. Le wiki a été inventé par Ward Cunningham en 1995, pour une section d'un site sur la programmation informatique qu'il a appelée WikiWikiWeb.

Le mot « wiki » vient du terme hawaïen wiki wiki, qui signifie « rapide » ou « informel ». Au milieu des années 2000 les wikis ont atteint un bon niveau de maturité et sont associés au Web 2.0. Créée en 2001, l'encyclopédie Wikipédia est devenue le wiki le plus visité au monde.



WIKIPÉDIA  
L'encyclopédie libre

- Accueil
- Portails thématiques
- Index alphabétique
- Un article au hasard
- Contacteur Wikipédia

- Contribuer
- Aide
- Communauté
- Modifications récentes
- Accueil des nouveaux arrivants
- Faire un don

- Imprimer / exporter
- Boîte à outils

- Autres langues
- Català
- Česky
- Deutsch
- English
- Español
- Suomi
- Magyar
- Italiano
- 日本語
- 한국어

Accueil Discussion

Lire Voir la source Afficher l'historique

Rechercher

## Bienvenue sur Wikipédia

Le projet d'encyclopédie libre que vous pouvez améliorer

1 097 192 articles en français

Version pour appareil mobile

Culture — Géographie — Histoire — Sciences — Société — Technologies

Liste des portails thématiques

### ★ Lumière sur

La **Macreuse noire** (*Melanitta nigra*) est une espèce de canard plongeur marin de la famille des anatidés. Le dimorphisme sexuel est prononcé chez cet oiseau, le mâle étant entièrement noir, avec une tache jaune orangé sur le bec, la femelle et les jeunes sont brun clair, avec les côtés du cou et de la tête plus pâles. La Macreuse noire habite le Nord de l'Europe et de l'Asie ; les populations américaines ont évolué vers une espèce désormais considérée comme séparée, appelée Macreuse à bec jaune (*Melanitta americana*).



La Macreuse noire hiverne en bandes et revient en mai sur ses aires de nidification où elle se reproduit sur les lacs, dans les roselières ou dans la toundra. La ponte compte cinq à huit œufs, couvés par la seule femelle, qui s'occupe également seule de ses canetons durant plusieurs semaines après leur naissance. La Macreuse noire plonge pour se nourrir. Son alimentation comprend principalement divers invertébrés où sont prédominants crustacés et mollusques, ainsi que vers et insectes, régime parfois complété à l'aide de petites pousses végétales ou de plantes aquatiques...

★ Lire la suite

Le **Léopard de mer** (*Hydrurga leptonyx*) est un mammifère carnivore de la famille des phocidés et représente actuellement la seule espèce du genre *Hydrurga*. Décrit pour la première fois en 1820 par le zoologiste et anatomiste français de Blainville, il fait partie des cinq espèces de phoques présentes en Antarctique.



Puissant et imposant — la femelle peut mesurer 4,5 mètres pour une masse de 600 kilogrammes — il doit son nom de léopard aux petites taches sombres qui

### W Présentation

Wikipédia est un projet d'encyclopédie collective établie sur Internet, universelle, multilingue et fonctionnant sur le principe du wiki. Wikipédia a pour objectif d'offrir un contenu librement réutilisable, objectif et vérifiable, que chacun peut modifier et améliorer.

Le cadre du projet est défini par des principes fondateurs. Son contenu est sous licence Creative Commons by-sa et peut être copié et réutilisé sous la même licence – même à des fins commerciales – sous réserve d'en respecter les conditions.

Actuellement, Wikipédia en français compte plus de deux mille articles distingués comme « articles de qualité » ou comme « bons articles ».

À propos de Wikipédia · Guide sur Wikipédia

### ✎ Participation

Chacun peut publier immédiatement du contenu en ligne, à condition de respecter les règles essentielles établies par la communauté ; par exemple, la

## **Edition**

L'accès en lecture comme en écriture à un wiki se fait avec un navigateur Web classique.

Les pages peuvent être visualisées dans deux modes différents : le mode lecture est le mode par défaut, le mode d'édition présente la page sous une forme permettant sa modification.

En mode d'édition, le texte de la page apparaît dans un formulaire Web. Le texte éditable est enrichi de quelques caractères respectant une syntaxe informatique, c'est le wikitext. Il permet d'indiquer la mise en forme du texte, de créer des liens, de disposer des images. Cette syntaxe est conçue pour être facilement assimilable et se distingue du langage HTML généralement utilisé pour rédiger des pages Web.

## **Administration des modifications**

Un wiki archive chaque modification avec sa date et son auteur. Il est ainsi possible de

revoir l'évolution de toute page et de revenir très simplement à une ancienne version. Il est également possible de revoir les contributions de tout visiteur. Il est enfin possible de suivre en temps réel toutes les modifications faites dans le site. Toutes ces

informations sont généralement archivées dans une base de données.

Ces pages de suivi ne sont évidemment pas librement modifiables. Il existe d'autres

pages administratives qui ne sont pas modifiables, comme les pages d'analyse des liens et les pages d'interface avec le système.

Fin ...