

# Les trois piliers de référencement sur le Web

Gheysen Jérémie

17 février 2016

## 1 Introduction

Conférence présidée par le Dr Ir Robert VISEUR (UMONS/CETIC). Travail au CETIC et est assistant dans le service de Management de l'Innovation Technologique à la FPMS à l'Umons.

## 2 Pilier de référencement naturel

Consiste à faire connaître les pages web d'un moteur de recherche. Lien commerciaux au dessus des liens référencés naturellement. Le référencement est très importants, les 3 premiers résultats sont les plus visités. Exposé axé principalement sur Google qui domine le marché avec plus de 90% des parts de marché en Europe. Moins de dominance au niveau de la Chine, Corée du Sud, Russie... (< 35%). La recherche Web ne se limite pas à la ressource textuelle : images, vidéos, flux RSS... Formats de données sur internet ?

- Pages web : Format standard de présentation ("HTML") (simple à référencer)
- Applications Web : Flash(binaire), Ajax(fb fait en JavaScript), ce sont des pages dynamiques, utilisation du doublage statique pour permettre le référencement.

Moteur de recherche c'est quoi ?

Automate en 3 parties :

- Un crawler (robot) : parcourt un ensemble de pages et les suit les liens de ces pages (pb avec les sites en Flash et Javascripts qui ne seront pas explorés par le robots). Il alimente un cache local.
- Indexeur : prend le contenu de la page mis en cache par le moteur. Analyse le texte, récupère les mots, les associe aux différentes pages. Principe du dictionnaire inversé (association "token" (mot) / page).
- Une interface d'interrogation.

Associer des actualités au site, ajouter des images liées à ces actualités vont aider à rendre le site visible dans Google WebSearch. (Recherche universelle, en dessous de la zone chaude (3 premiers liens)).

### Contrôler l'indexation

- Signaler un nouveau site en ligne via le formulaire d'ajout Google
- Ajouter un siteMap, plan du site en (XML à vérifier) pour permettre au robot de mieux visiter le site (pas forcément nécessaire).
- Configuration du fichier <robot.txt> = fichier (et balise META) permettant d'indiquer au robot ce qu'il peut faire ou ne pas faire sur le site. Permet de spécifier aux robots des moteurs de rechercher les zones qu'il peut indexer ou pas dans le site. **Cache Google : permet de récupérer des pages en plus ou moins bonne forme d'un site down**

Lors de la création d'hyperliens, écrire sur l'hyperlien un texte ayant un intérêt en terme de référencement.

Utilisation les balises les plus importantes pour mettre en évidence les expressions (mots-clés) dans le texte du site.

Exploiter les liens externes et internes :

- Page boostée par les citations, plus il est cité, plus il est important, avec un poids plus important si le lien provient d'un site de poids plus important !

- Popularité différente des pages indexées (en fonction du nb de lien (et du PR des liens).
- PR(PageRank) d'une page (cote de la page internet), plus le PR était élevé plus le site était populaire. Avant ne prenait en compte que le nb de lien (pb de triche avec des générateurs), maintenant prend en compte beaucoup plus d'éléments comme la fréquence de mise à jour, ... A évolué vers le TrustRank(TK).

Surveiller la qualité des liens entrants : Popularité, cohérence thématique/géographique/linguistique,

..

Attention à la duplication des liens, achat de plusieurs noms de domaine et redirection statique vers un seul => Google n'apprécie pas. Trop de contenu générique sur une page => éviter! Réécriture d'URL a été mal implémentée => empêcher la coexistence entre URL réécrites et URL non réécrites (par ex : `www.domaine.com` et `domaine.com`) || Eviter le `index.php` si jamais changement de langage.

Suivre l'audience d'un site : des outils statistiques le permettent, également mesure de qualité de référencement. Trois types de systèmes de statistiques :

- Basés sur les logs du serveurs (AWstats) : log toutes les connexions mais richesse moyenne
- Applications hébergées utilisant un tag JavaScript qui ping un serveur Piwik et enregistre l'entrée de la connexion (pas faisable si bcp de connexions sur le site) + permet de localiser les points chauds du site
- Le plus utilisé : Google Analytics, très riche fonctionnellement !

Recherche universelle : les résultats de Google Web sont mixés avec les vidéos (Youtube), carte (Maps), images (Google image), actualités (Google Actualités) => Nouveaux angles d'attaques pour être bien référencé!

### 3 Pilier de référencement commercial

Publicités encodées par les annonceurs dans Google, ce dernier assure une cohérence contextuelle. Paiement au clic, valeur du prix aussi déterminée par la concurrence liée mot clé. Requêtes spécifiques sont peu chères car peu recherchées. On peut aussi choisir une zone géographique. Google permet de distinguer le volume de visites liées aux publicités, permet d'analyser la rentabilité du site. Le fait de faire de la pub ne permet pas d'avoir un meilleur référencement naturel, les deux notions sont totalement indépendantes.

### 4 Pilier de référencement social

Idée d'avoir une visibilité dans les réseaux sociaux. Avoir une visibilité dans LinkedIn, tant pour l'entreprise que les employés de l'entreprise. Avoir une présence dans les outils comme Facebook devient de plus en plus intéressant. Les réseaux sociaux permettent également de créer des liens vers des sites qui parlent de l'entreprise. Cela peut générer des visites sur notre site.