

# Chronique

## de la recherche sur le web

par Marc Duval

19 août 2003, vol. 3, n° 7

### Référencement: Deux études de cas

**P**our les uns, les entreprises commerciales sont le fléau de la recherche, pour les autres, les automates de recherche sont une bénédiction. Elles sont non seulement dans les résultats de recherche mais elles les entourent. Elles ne sont plus tapageuses mais d'une discrétion envahissante que la Federal Trade Commission a essayée de policer.

Connaître les techniques de référencement, c'est savoir pourquoi une page est affichée parmi les dix premières même si elle ne contient aucun mot d'une requête en apparence. C'est aussi une façon de jauger la qualité d'un automate de recherche.

Trois techniques de référencement sont visitées. Ce sont celles de la balise titre, du réseautage et du texte invisible. L'utilisation de la balise titre et du texte invisible est bien connue dans le milieu du référencement. C'est une de leurs variantes qui est montrée. Deux entreprises sont étudiées pour illustrer les trois techniques: Canam Steel et Boa-Franc.

Canam Steel est UNE entreprise manufacturière de produits d'acier, filiale de Canam Manac, une grande entreprise du Québec. Chaque filiale a son propre nom de domaine et son propre site. Boa-Franc est une moyenne entreprise tout à fait respectable comme il y a en tant.

#### La balise Titre

La balise Titre est surtout connue sous le nom anglais « meta-title ». Elle est le titre qui apparaît dans un résultat de recherche et dans la barre supérieure d'un navigateur. Elle est reconnue comme un élément fondamental. Elle est la carte de visite d'une entreprise. Elle contribue à son

image.

En 2002, le titre de la page d'accueil de Canam Steel était « [Canam Steel, steels joists, steel components](#) ». Ce titre ne paie pas de mine. C'est une succession de mots qui ne donne pas la vraie nature de l'entreprise. Le marché visé est les États-Unis. Les mots retenus par la société de référencement pour l'optimisation sont : joist, joists, steel joist, steel joists, metal joist, metal joists, joist manufacturer, joists manufacturer joist manufacturing et joists manufacturing La technique consiste à utiliser le style télégraphique aussi original qu'une liste d'épicerie. Le titre devient : Canam Steel (**Steel Joist Manufacturer**) -- **steel joist.steel joists.steel building, metal joists.joist manufacturing**. La première partie est convenable (Canam Steel (**Steel Joist Manufacturer**)). La seconde l'est moins. Chacun de ces termes n'est pas repris dans la page pour conduire l'internaute vers l'un des produits proposés. Ce style télégraphique est efficace paraît-il et couramment pratiqué.

#### Le réseautage

Le [classement de Google](#) s'appuie sur l'analyse des hyperliens et le réseau qu'ils tissent. Les référenceurs en ont pris acte en exploitant cette caractéristique avec plus ou moins de bonheur. Le classement dépend du nombre de liens externes pointant vers le site visé. C'est une variante de la technique du site miroir. Le site miroir est la copie d'un autre site mais avec une adresse différente. Le réseautage de Canam Steel est particulier. Il est composé d'au moins six domaines : canam.cc, steelplus, canammanac, spnkolkata, acierplus et canamjoist.com. De ces six noms, canamjoist est introuvable. Les cinq autres réfèrent à des entreprises différentes mais avec le même titre ou une variante. A l'intérieur de chacun d'un, une page fait appel au même fichier de Canam Steel à partir d'une base de données Lotus. En utilisant l'une des requêtes retenues (joist, joists, steel joist, steel

joists, metal joist, metal joists, joist manufacturer, joists manufacturer joist manufacturing et joists manufacturing), l'un des cinq sites apparaît et conduit à la page d'accueil de Canam Steel. Le site qui revient le plus souvent est celui de canam.cc. Voici les adresses de Canam Steel émanant des cinq sites :

- <http://www.canam.cc/CANAM.NSF/Francais!OpenPage>
- <http://www.steelplus.com/canaman.nsf/>
- <http://www.canammanac.com/canaman.nsf/>
- <http://www.spnkolkata.com/canaman.nsf/>
- <http://www.acierplus.com/canaman.nsf/pageacceuil/0>

Le résultat de ce travail est mitigé.

Contrairement à ce que l'on pourrait penser, le site n'a pas été optimisé pour Google mais pour AlltheWeb/Lycos. Canam Steel est bien positionné sur AlltheWeb/Lycos avec un pointage de 9 requêtes sur 10 donnant une page du site en premier grâce à son réseautage. L'entreprise réussit un pointage de 5 sur 10 sur AltaVista, tout comme sur Teoma ; Google obtient un pointage de 4 sur 10 et Inktomi/MSN, 2 sur 10. Le tandem Yahoo!/Google n'a pas fait mieux, le site n'étant pas référencé dans Yahoo!

Voici le tableau de positionnement de la page d'accueil de Canam Steel. L'astérisque suivant le nombre indique qu'il s'agit d'un page du site du réseau menant à la page d'accueil de Canam Steel.

#### Position de la page d'accueil de Canam Steel au 11 août 2003

Sites de recherche	Joist	Joists	Steel joist	Steel joists
AlltheWeb	16	7	2	2
AltaVista	22*	17*	10*	7*
Google	102	62	4	2
MSN	77	69	46	68
Teoma	92*	108	17	2
Yahoo/Google	83	55	3	2

Sites de recherche	Metal joist	Metal joists	Joist manufacturer	Joists manufacturer
AlltheWeb	7	8	2	2
AltaVista	8*	12*	7*	3*
Google	16* / 80	25	1	4
MSN	300+	9	129*+	4
Teoma	11*/14	14	1	8
Yahoo	15*/64	19	1	3

Sites de recherche	Joist manufacturing	Joists manufacturing
AlltheWeb	9	3
AltaVista	7*	7*
Google	18	33
MSN	31	6
Teoma	1	1
Yahoo!	12	26

### La technique du texte invisible

La troisième technique est celle du texte invisible. Les techniques les plus utilisées consistent à cacher des URLs ou une liste de mots soit avec des petits caractères, soit avec la même couleur pour les lettres et le fond. Ces techniques servent à tromper l'automate de recherche. Elles ont été reconnues et mises à l'index. Les automates ont intégré des algorithmes afin de les détecter.

Le site de Boa-Franc est un cas tout aussi intéressant que Canam Steel. L'entreprise est spécialisée dans les planchers de bois franc. Les référenceurs ont utilisé une technique de texte invisible controversée. Elle consiste à employer la balise `<noscript>` qui rend invisible les phrases au profane dans la page de code source. Elle est censée être lue et acceptée comme élément de recherche par les automates. La conception de cette partie de la page prend deux critères de classement important : l'en-tête de titre 1 et les hyperliens.

#### L'en-tête de titre

Les en-têtes de titre ont plusieurs niveaux : l'en-tête 1 est le premier niveau, celui du titre la page, puis suit le titre de chapitre - en-tête 2, le titre de paragraphe - en-tête 3, jusqu'à l'en-tête 6. Un mot placé dans un en-tête de titre 1 pourrait aider à classer une page au-dessus de celle dont le mot est dans le texte.

La balise `<noscript>` se trouve sur la page en anglais de la page d'accueil de Boa-

Franc. Voici le début de la section réservée à cette balise.

```
<noscript> <h1>Mirage Hardwood Floors
&mdash; Hardwood Flooring and Wood Floors
Products</h1> <h3>Hard wood flooring and
hard wood floors</h3> <h1>Mirage
Floors</h1> <a
href="http://www.boa-franc.com/boaeng/plan_mi
r/pourquoi1.asp">Why
choosing Mirage hardwood floors or hardwood
flooring or wood floors ?</a> </noscript>
```

Voici comment apparaissent les mots mis en en-tête 1 :

## MIRAGE HARDWOOD

Ces mots sont visibles lorsque les codes `<noscript>` et `</noscript>` sont retirés de la page source.

#### Les hyperliens

Les automates sont friands des hyperliens. C'est pourquoi tous les liens internes ont été ajoutés comme ceci :

```
<a href="http://www.boa-franc.com/boaeng/
plan_mir/collection1.asp">Mirage
hardwood floors, hardwood flooring and wood
floors Collections</a>
```

Le résultat de cette combinaison de méthodes d'écriture met la page de Boa-Franc avec les mots « [mirage hardwood](#) » en première position sur Google. Elle est première non

seulement parce que ces mots sont dans le titre et dans la page visible, mais aussi parce qu'ils sont dans un en-tête de titre 1, hyperliés et répétés, le tout invisible à l'internaute. Ajoutons l'inscription du site dans le répertoire de l'ODP. En contrepartie, la page qui n'a pas cette « optimisation » est « oubliée ». C'est le cas de la version en français de la page d'accueil titrée « Plancher bois-franc et plancher de bois franc par Mirage ». Malgré le lien sur la page d'accueil en anglais, cette page n'a pas été indexée par Google. La technique de proximité des mots (répétition de plancher et de bois-franc avec et sans la préposition ) aura été inutile.

### Conclusion

Plusieurs référenceurs usent de méthodes parfois à la limite du bon goût en utilisant un style télégraphique pour les titres, parfois à la limite de l'indélicatesse avec le détournement du réseautage et de la rédaction d'une page web. Ils

arrivent souvent à leurs fins avec une augmentation de trafic au détriment de l'image de l'entreprise. Les entreprises insouciantes n'ont-elles pas vendu leur âme au diable?

---

#### Référence:

Google technology inc. (Page consultée le 19 août 2003). Google - Informations pour webmestres. [En ligne]. Google, 2002. <http://www.google.fr/intl/fr/webmasters/>

---

Marc Duval  
Bibliothécaire professionnel  
Service de recherche documentaire DSI  
1280-4, bd de Montarville.  
Longueuil. Québec. Canada  
J4B 8B4  
[marcduval@dsi-info.ca](mailto:marcduval@dsi-info.ca)  
<http://www.dsi-info.ca>

---

Mots clés : Bibliothèques (vedettes matières) : recherche sur Internet ; recherche de l'information ; recherche documentaire ; ouvrages de références ; gestion de l'information / Nomade : recherche sur le Web ; moteurs de recherche / ODP : recherche sur le Net ; moteurs de recherche / Toile du Québec : gestion de l'information ; Internet - Guides ; guides perfectionnement à la recherche ; moteurs de recherche / Voila : internet - recherche d'info sur le web ; moteurs de recherche / Yahoo : recherche sur le Net ; recherche sur le Web ; recherche et traitement de l'information ; internet - information et documentation ; moteurs de recherche / Pages jaunes : renseignements-service Longueuil ; Identificateur: Référencement