

---

numéro trente quatre ( #34 )

le 12 mars 2002

---

---

## Sommaire

---

- [Sommaire](#)
- [Editorial](#)
- [Enfin, les informations](#)
- [Les outils de recherche francophones](#)
- [Entretien avec Philippe Préget](#)  
de Mirago.
- [Conclusion](#)

[Le dernier numéro de Moteurzine](#)

[Le dernier Moteurzine en PDF](#)

[Les archives de Moteurzine](#)

[Ecrire à Moteurzine](#)



---

## Editorial

par [CHRIS HEDE](#)

---

### un édito simple et court

<mode autopromotion on>

*Les indices de popularité n'ont plus de secret pour vous, vous savez jouer avec les pages satellites et les pages alias, vous avez des petits (voire des gros) secrets de référenceurs à dévoiler... c'est le moment ou jamais de les partager avec les lecteurs de cette newsletter. Si vous vous sentez une âme de rédacteur ou de journaliste, il est possible de mettre des articles en alternance avec les entretiens sur « Moteurzine ».*

De même vous êtes une des nombreuses entreprises de référencement francophones (ou une webagency avec un pôle référencement) et vous ne savez pas comment vous démarquer auprès de vos clients par rapport à vos concurrents ? Associez donc sérieusement votre nom à celui de « Moteurzine » et soyez LA société de référencement présente sur « le bulletin des spécialistes des outils de recherche francophones ».

<mode autopromotion off>

N'hésitez pas à prendre contact avec moi pour l'un ou l'autre de ces points. Merci.

Bonne lecture à tous et à bientôt.

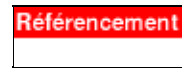
PS : Ce numéro de Moteurzine est [disponible, pour la première fois, en PDF](#), si vous trouvez utile d'avoir cette newsletter sous ce format, n'hésitez pas à me le signaler.

---

## Espace publicitaire

---

Améliorez votre **indice de popularité** en utilisant **Refgratuit.com**. Vous pouvez référencer votre site dans une vingtaine d'outils de recherche francophones (moteurs et annuaires) : Abacho, Adminet, Alltheweb, Antisearch, Google, Gotcha, Hotbot, Janas, Kagibi, Lagitane, Maximini, Netscan, Planète virtuelle, Tritou, Voila et Voiloo.



---

## Enfin, les informations

---

[Diffusez gratuitement votre actualité liée au monde des outils de recherche et du référencement.](#)

**[09/03/2002 : "Euroseek, un retour prévu pour le premier trimestre 2002"](#)**

Euroseek (la société) est d'origine suédoise et a été déclarée en faillite en septembre 2001. Euroseek (le site) a été racheté 26000 euros par la société suédoise Wordlight. Wordlight a annoncé le retour d'Euroseek pour le premier trimestre 2002 et espère récupérer une partie des 80 millions de visiteurs uniques mensuel.

**[09/03/2002 : "Looksmart UK se converti à Google"](#)**

Quand l'annuaire Looksmart UK n'a pas de réponses suffisantes et satisfaisantes, la recherche s'effectue normalement sur Inktomi. Depuis peu, elle se fait maintenant sur Google. Quand ce changement va être généralisé à Looksmart France ? Et à Looksmart US ?! Une question de semaines, voire de jours selon certains bruits.

**[08/03/2002 : "L'annuaire de Scoot France sur minitel"](#)**

L'annuaire de professionnels Scoot France, qui était disponible sur le Web et par téléphone, sera désormais accessible par minitel 36 15 SCOOT.

**[06/03/2002 : "Yahoo.fr s'ouvre aux femmes en attendant le Mondial"](#)**

On dit souvent que les femmes sont l'avenir du Web et Yahoo France ne veut pas rater le coche. Le portail va prochainement ouvrir un espace intitulé "Bien-être" qui proposera des informations et des services destinés en priorité à une population d'internautes féminins. Pour se faire, des partenariats ont déjà été signés avec des éditeurs de contenu comme les magazines Top Santé, Psychologie, Elle et Vogue. En préparant ce nouvel espace féminin, Yahoo suit le mouvement initié par la plupart des grands portails. AOL France a ainsi ouvert une chaîne féminine sur son

service abonnés. MSN France, Lycos France et, plus récemment, Voila.fr se sont dotés à leur tour d'espaces dédiés aux femmes. Sur ce créneau, TF1.fr s'est singularisé en optant cette fois pour une marque dédiée : Plurielles.

[05/03/2002 : "Les moteurs de recherche se font une place dans le milieu des entreprises"](#)

Les dernières annonces des principaux outils de recherche (Google et Altavista) montrent que les moteurs envisagent une croissance de leurs revenus en provenance du marché des entreprises. Le 11 février 2002, Google est entré chez les entreprises en lançant le "Google Search Appliance" une solution intégrée matériel / logiciel pour les intranets institutionnels et les servers web. Puis, le 23 février 2002, AltaVista Software a élargit la gamme de ses produits de recherche pour les entreprises en sortant le "AltaVista Desktop Search".

[04/03/2002 : "Google suit l'actualité à travers son logo"](#)

Tout le monde a remarqué avec le Nouvel An et les Jeux Olympiques que Google change son logo quotidiennement pour coller à l'actualité et surtout, selon moi, apporter plus de fraîcheur au moteur de recherche ainsi qu'un côté plus convivial. L'équipe d'infographistes derrière tous ces logos est composée d'une seule personne : Dennis Hwang.

[02/03/2002 : "Espotting, le réseau de "pay-per-click" numéro un en Europe"](#)

Selon une étude menée par Nielsen/NetRatings, Espotting propose en Allemagne 31% de plus de recherche que son principal concurrent, Overture.

---

## Espace publicitaire

---



---

## Les outils de recherche francophones

---

### 09 Mars 2002 :

TOP | internet | guide généraliste | s

[Samaritania](#) : Ressources et services gratuits pour webmasters et surfeurs.

TOP | internet | guide généraliste | c

[Colletoizy](#) : Le référenceur de sites internet.

TOP | régional | europe | france | auvergne

[Auvergne-net](#) : Le portail des sites internet de l'Auvergne.

### 08 Mars 2002 :

TOP | régional | europe

[Beskid](#) : Le premier annuaire francophone consacré à la Pologne.

TOP | régional | europe | russie

[Moscoupolis](#) : Annuaire des sites consacrés à la Russie.

TOP | internet | guide généraliste | a

[Annuaire-net](#) : Un éventail des meilleurs sites gratuits du Web classés par thèmes.

**07 Mars 2002 :**

TOP | régional | europe

[Roumanie-mania](#) : Un maximum de liens vers la Roumanie.

TOP | régional | europe | belgique

[Pim](#) : Découvrir la Belgique et y vivre.

TOP | santé et spiritualité | religion | christianisme

[Custodia](#) : Des chrétiens sur le web...

TOP | institutions | enseignement

[Zool2000](#) : Ressources pédagogiques pour l'école primaire.

TOP | régional | europe | suisse

[Kouik.ch](#) : Moteur de recherche / annuaire de sites entièrement consacrés à la Suisse romande.

**06 Mars 2002 :**

TOP | commerce et économie | travail | recrutement

[Tchatchemploi](#) : Des forums gratuits pour déposer des offres d'emploi et des CV.

**05 Mars 2002 :**

TOP | internet | guide généraliste | t

[Tysami](#) : Annuaire de recherche pour trouver les sites que vous aimez !!!

TOP | loisirs et sports | vedette

[Voistars](#) : Annuaire dédié aux célébrités.

TOP | régional | asie

[Netarménie](#) : Annuaire des sites de l'Arménie.

TOP | régional | europe | russie

[Moskovia](#) : Annuaire des sites russes et russophiles.

TOP | loisirs et sports | sport | rugby

[Francerugby](#) : Annuaire de Francerugby.

[Plus2rugby](#) : Le guide du rugby sur Internet !

TOP | commerce et économie | travail | recrutement

[ABannonces.com](#) : Emploi en informatique et télécoms.

[Alinea](#) : Emploi du milieu de la santé (médical et paramédical).

TOP | commerce et économie | immobilier

[Coolimmo](#) : Portail immobilier francophone de diffusion d'annonces de location et vente.

**04 Mars 2002 :**

TOP | loisirs et sports | sport | football

[Le kop](#) : Guide–annuaire des sites de supporters.

**02 Mars 2002 :**

TOP | régional | afrique | algérie

[Algeria search](#) : Le moteur de recherche du web algérien.

TOP | sexe | photographie

[Photoscope](#) : Annuaire des meilleurs sites web de photographie.

TOP | internet | guide généraliste | o

[Owel](#) : Annuaire des sites web d'Owel.

TOP | commerce et économie | annonces

[Service de diffusion d'annonces](#) : Site permettant de diffuser son annonce sur de nombreux sites de petites annonces à la fois.

TOP | art et culture | audiovisuel | musique

[Trocmusic](#) : Annuaire professionnel de la musique et de l'audio.

**01 Mars 2002 :**

TOP | internet | guide généraliste | a

[Annuaire 555](#) : Annuaire des sites internet francophones.

TOP | loisirs et sports | sport | hippisme

[Clamarcourses](#) : Annuaire des sites liés au hippisme et pronostics.

TOP | loisirs et sports | jeux | jeux de simulation

[Jdr central](#) : Portail des ressources pour maître du jeu.

[Le python](#) : Les anneaux du jeu de rôle.

TOP | régional | autres pays

[Best of Brésil](#) : Sélection des meilleurs sites sur le Brésil.

---

## Espace publicitaire

---



Refgratuit.com vous permet d'avoir votre propre outil de référencement aux couleurs de votre site. Avec ce service, vous pouvez ainsi capter de nouveaux visiteurs et les garder plus longtemps chez vous.

Pour voir Refgratuit.com aux couleurs d'Enfin :

<http://www.enfin.refgratuit.com/>.

[Contactez nous](#) pour en savoir plus !!

---

## Entretien avec...

<http://www.mirago.fr/>

Philippe Préget, « Head of Operations » à Mirago

---

«**Moteurzine**». -- Bonjour. Tout d'abord, pouvez-vous vous présenter à nos lecteurs et nous expliquer comment vous êtes arrivé à travailler sur Internet et pour Mirago ?

**Philippe Préget.** -- Je m'appelle Philippe Préget. Je suis « Head of Operations » à Mirago. J'ai travaillé pendant plusieurs années dans le développement de bases de données avant d'être impliqué dans le projet Mirago. Comme je suis français et que j'utilise régulièrement le web français, j'ai attaché un soin particulier à faire du service français un outil efficace.

«**Moteurzine**». -- Pouvez-vous nous raconter l'histoire de Mirago, qui est, si je ne me trompe pas, une société présente sur Internet depuis au moins 1997 ?

**Philippe Préget.** -- Le développement de Mirago a débuté au sein de la société Compssoft, une compagnie britannique spécialisée dans les logiciels de base de données. Compssoft crée et distribue ses propres solutions depuis 1979. Au coeur de tout système de base de données se trouve un moteur d'indexation qui enregistre et gère une grande variété de données. C'est cette technologie qui est à l'origine du projet Mirago.

En effet en 1996, Compssoft lance un projet visant à appliquer ses techniques d'indexation à de larges sources de données, notamment Internet. A la suite du projet initial de RD au sein de Compssoft, Mirago est créé et l'équipe en charge du projet intègre la nouvelle entité.

Un service-test est lancé à la fin 1998, puis le service complet en 1999. En 2000, Mirago lève suffisamment de fonds pour développer et déployer sa stratégie de recherche régionale.

Aujourd'hui Mirago exécute près de 1,5 millions de requêtes par jour pour la Grande-Bretagne (GB) uniquement et prépare actuellement des projets pour la France, l'Allemagne et l'Espagne.

Mirago est une compagnie privée, comptant parmi ses actionnaires l'un des plus

grands groupe de presse régionale de la GB.

«**Moteurzine**». — Comment pouvez-vous définir exactement Mirago ?

**Philippe Préget.** — Avec l'accroissement continu de données disponibles sur le net, il devient de plus en plus nécessaire de mettre en place des « filtres régionaux » afin d'éliminer les données non géographiquement pertinentes pour les utilisateurs de moteurs de recherche. Notre technique de sous-catégories géographiques à l'intérieur même des index générés par nos robots répond à cette nécessité. Nous nous décrivons donc comme un moteur de recherche national offrant des recherches régionales.

«**Moteurzine**». — La situation des annuaires et moteurs de recherche en Angleterre est souvent confondue avec celle des outils à l'international. Vous pouvez nous dire quels sont les principaux sites de recherche en Angleterre ?

**Philippe Préget.** — Les principaux sites de recherche présents en GB sont principalement des sites américains offrant leur service de recherche via un site en .uk. Il y a également des méta-moteurs tels que Ask Jeeves. Nous pensons que nous disposons du plus grand index britannique avec environ 50 millions de pages.

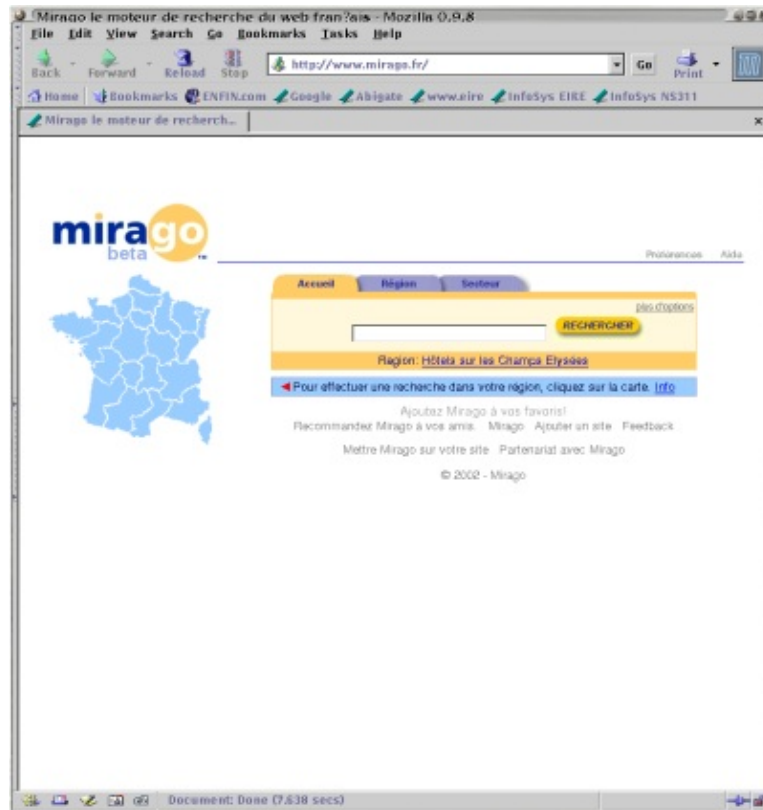
«**Moteurzine**». — Quels sont vos principaux concurrents au niveau national, européen comme international ?

**Philippe Préget.** — Nous n'avons pas encore vu d'autres moteurs de recherche adopter une optique régionale de l'Internet pour un pays donné. Je suppose que nous sommes en concurrence en général avec les autres moteurs de recherche et annuaires, bien que nous ayons notre propre angle.

«**Moteurzine**». — Mirago, société anglaise, vient d'ouvrir une version française de son outil de recherche. Quel est votre objectif ?

**Philippe Préget.** — Notre objectif à long terme est d'accroître le nombre de nos utilisateurs en France. Afin de parvenir à cet objectif, nous envisageons de développer des partenariats avec des sociétés françaises capables d'intégrer nos index dans leur contenu.

Notre objectif à court terme est de nous concentrer sur la qualité et l'exhaustivité de notre index, afin d'offrir la meilleure expérience utilisateur.



*Copie écran de la page de garde du site Mirago*

«**Moteurzine**». — Que pensez-vous de la situation du marché des outils de recherche en France ? Quel va être votre positionnement sur ce marché ?

**Philippe Préget.** — Il semble que certains moteurs de recherche ont une vision un peu simpliste de l'Internet français qu'ils réduisent aux seuls domaines en .fr. Notre index comprend tous les sites en .org, .com, .net qui sont basés en France en plus des sites en .fr. De plus nous avons constaté qu'il y a très peu de services de recherche régionale.

De la même façon que l'Internet global a été « nationalisé » pour faire face aux énormes quantités de données sans pertinences géographiques présentées aux internautes, une vision régionalisée de chaque pays est devenue un impératif pour permettre aux internautes de localiser les informations, produits ou services pertinents pour la communauté dans laquelle ils vivent et travaillent. Mirago offre ce service.

«**Moteurzine**». — Comptez-vous ouvrir d'autres versions de votre outil dans d'autres pays européens comme l'Italie ou l'Allemagne ?

**Philippe Préget.** — Oui. Nous sommes en phase beta pour l'Espagne et venons de lancer un robot pour l'Allemagne.

«**Moteurzine**». — Comment se gère cette version française au niveau humain mais aussi au niveau technique ?



**Philippe Préget.** — L'aspect « utilisateur » est géré par des nationaux français basés en GB. Notre équipe centrale en charge des opérations s'occupe également des opérations techniques pour la France.

«**Moteurzone**». — Quelle est la technologie logiciel et matériel derrière Mirago ? Tous vos serveurs sont en Angleterre ?

**Philippe Préget.** — Tous les logiciels utilisés pour constituer et interroger nos index, fournir du contenu et montrer des liens sponsorisés sont des technologies propriétaires (à part bien sûr nos serveurs web IIS !). Nous assemblons et maintenons notre propre hardware.

Nos liens avec l'Internet français montrent des temps de réponse très bas, de ce fait il ne nous est pas nécessaire d'être localisés en France.

«**Moteurzone**». — Avez-vous une base de données différentes pour chaque pays ? Et pourquoi ce choix ?

**Philippe Préget.** — Une base de données propriétaire est utilisée pour contrôler les robots et classer par région et secteur les sites indexés par ces derniers. Il y a une base de données de contrôle et un cluster robot par pays. Cette approche par pays nous permet d'ajuster la durée des cycles d'indexation pour chaque pays, plutôt que d'être contraint par un long robot général.

«**Moteurzone**». — Comment un webmaster (ou une société de référencement) doit faire pour référencer son site sur Mirago ?

**Philippe Préget.** — Il y a un lien vers la page « Ajouter votre site » sur la page d'accueil de notre site.

«**Moteurzone**». — Est-ce qu'il y a certaines choses à faire ou ne pas faire pour optimiser ce référencement ?

**Philippe Préget.** — Nos algorithmes de positionnement ont pour objectif de détecter les thèmes narratifs d'une page et de pénaliser une utilisation non naturelle des textes ou une fréquence anormale d'un mot. Le meilleur conseil est de ne pas utiliser de spam et d'utiliser judicieusement les outils de design telles que les balises de description pour fournir des résumés de résultats ad hoc.

Ce qui s'est produit en GB c'est que plutôt que d'essayer de deviner les algorithmes des moteurs de recherche, les sociétés de référencement placent à présent pour leurs clients des liens sponsorisés effectifs.

«**Moteurzone**». — En regardant le fichier log de mon site, j'ai découvert :  
**217.33.60.210 - - (...) "www.mirago.fr/scripts/frinfo.asp"**  
**"PierreLeRobotMirago"**. Pierre, c'est le nom de votre robot ? A quelle profondeur navigue-t-il dans un site ?

**Philippe Préget.** — Oui c'est nous ! Actuellement il va à un niveau équivalent au chemin le plus court via 6 clics de navigateur à partir de la page d'accueil. Nous

appelons cela les « hop count ». A présent que nous avons une idée du nombre de pages que comprend le web français, nous envisageons d'augmenter les « hop count ».

«**Moteurzone**». — Est-ce qu'il prend en compte les fichiers PDF ? Et comme Google, lit-il aussi les documents en .DOC ou .XLS ?

**Philippe Préget**. — Non, pas pour l'instant.

«**Moteurzone**». — Comment votre robot prend-il en compte ces différents éléments : l'url, le titre, les méta "description" et "keywords", le texte sous forme de commentaires, le texte dans une page, les liens vers d'autres sites, le lien d'autres sites vers la page ?

**Philippe Préget**. — Nous indexons tous ces éléments en leur accordant différents degrés de pertinence. Nous stockons à la fois les liens partant et pointant vers une page afin d'en déduire un index de données sur les sites populaires et de permettre aux webmestres de voir les liens pointant vers une page spécifique. Ces deux éléments sont disponibles en tant qu'options de recherche sur le site français.

«**Moteurzone**». — Combien avez-vous de sites et de pages dans votre base ?

**Philippe Préget**. — Le robot en cours a pour l'instant indexé 15 millions de pages sur approximativement 500.000 sites (.fr, .com, .net et .org). Sachant que nous sommes nouveaux sur le marché français, nous espérons que ce nombre croîtra rapidement au fur et à mesure que nous indexerons les premières soumissions.

Notre objectif est d'être le plus grand index pour France.

«**Moteurzone**». — Quels sont les choses que vous considérez comme du spam dans un site web ?

**Philippe Préget**. — Nous définissons spam comme tous les cas dans lesquels une page tente artificiellement d'accroître son importance au-delà de ce qu'un être humain lui aurait accordé. Les artifices tels que les textes cachés (caractères blancs sur fond blanc), les mots excessivement répétés et les caractères microscopiques sont autant d'exemples de spam. Bien sûr il existe beaucoup d'autres moyens utilisés par certains webmestres, que je ne mentionnerai pas afin de ne pas encourager l'utilisation de ces pratiques perfides!

A titre d'indications aux web développeurs et webmestres, je rappellerai que notre objectif est de faire en sorte que le logiciel d'indexation Mirago prenne des décisions similaires à celle d'un éditeur quant à l'importance des pages. En plus des textes cachés, notre logiciel cherche également tout ce qu'un éditeur utilisant un navigateur standard ne peut voir. Un exemple pour illustrer mes propos : il y a ce que nous appelons le « rapid client pull », il s'agit de pages utilisant des balises d'actualisation pour rediriger le navigateur vers une autre page dans un délai minimal. Nous considérons que le texte sur ce genre de page n'a pas été créé dans

le but d'être lu par un humain. En conséquence, nous traitons ces pages de la même manière !

Bien que vous ayez sans nul doute vu le robot Pierre lire des pages web dans les fichiers logs de votre serveur, vous aimeriez également considérer les autres « navigateurs » qui ont apparemment rapatrié des pages web. Les moteurs de recherche comme Mirago utilise une variété de techniques pour identifier le spam ou les pages cachées. Une des techniques utilisées consiste à lancer périodiquement un robot « déguisé » qui va lire quelques pages de votre site sélectionnées au hasard, mais avec une adresse IP différente. Les robots « déguisés » s'identifient comme étant une version récente d'un des navigateurs les plus communs. Les différences entre ce que peuvent lire le moteur original et le moteur « déguisé » sont utilisées pour détecter les sites qui tenteraient de biaiser l'indexation des pages web.

Le meilleur conseil que je puisse donner aux personnes développant des pages web est le suivant : généralement un texte bref et concis assure un bon positionnement. Evitez d'utiliser des artifices et n'oubliez pas que nous indexons les pages de millions de sites, donc nous avons probablement vu la plupart des méthodes de spamming couramment utilisées. Comme d'autres moteurs de recherche, nous favorisons les pages qui ne tentent pas de compromettre l'intégrité des index Mirago.

«**Moteurzine**». -- Dans un avenir plus ou moins proche, pensez-vous faire comme beaucoup d'autres outils en proposant un référencement payant dans Mirago ?

**Philippe Préget.** -- Nous y avons pensé mais nous n'avons pas de projet dans l'immédiat.

«**Moteurzine**». -- Avez-vous des projets de mettre des liens payants sur vos pages résultats ? Cela se fera plutôt avec la société anglaise Espotting, la société américaine Overture ou bien comme Google, avec votre propre solution ?

**Philippe Préget.** -- C'est le modèle que nous avons adopté en GB en utilisant notre propre solution. Les utilisateurs peuvent créer et gérer leurs comptes en ligne. Notre système actuel utilise un positionnement basé sur un modèle de coût par clic (via un mécanisme d'enchères), mais nous cherchons actuellement à introduire dans l'équation le ratio de « click through » de chaque listing. Une fois que nous aurons développé une base d'utilisateurs qui intéressera les annonceurs, nous lancerons le même modèle en France.

«**Moteurzine**». -- Quel est le modèle économique derrière Mirago ? La publicité et les bannières rapportant peu de revenus, comment pouvez-vous gérer votre développement ?

**Philippe Préget.** -- Notre modèle de revenu repose sur l'ajout de listings sponsorisés pertinents au-dessus de nos résultats de recherche. Ces sites sont clairement identifiables par l'utilisateur.

«**Moteurzine**». — Différentes crises ont secoué les entreprises Internet en 2001. Comment les a ressenties Mirago ?

**Philippe Préget.** — Nous avons développé un modèle de revenu viable et avons grandi naturellement plutôt qu'à travers d'extravagantes campagnes de marketing.

«**Moteurzine**». — Comment voyez-vous l'évolution des outils de recherche dans les mois à venir et plus particulièrement celle de Mirago ?

**Philippe Préget.** — Tôt ou tard, les autres moteurs de recherche vont devoir se poser la question de pertinence géographique à laquelle nous répondons. De plus, bien qu'il y ait une contrainte à générer des revenus, tous les moteurs ont encore besoin d'imposer un important contrôle éditorial sur les listings sponsorisés afin de préserver la pertinence des résultats.

«**Moteurzine**». — Pour finir, quel est votre site préféré et pourquoi ?

**Philippe Préget.** — [www.bobthebuilder.com](http://www.bobthebuilder.com) parce que ma fille de 4 ans peut y jouer en se prenant pour papa !

«**Moteurzine**». — Merci d'avoir pris le temps de répondre à toutes ces questions.

---

---

#### ATTENTION

La liste des abonnés n'est pas disponible. Nous ne la donnons ni ne la revendons à personne. En revanche, vous pouvez sponsoriser notre lettre d'information : contactez-nous pour discuter des modalités.

[Abonnement](#) gratuit à  
« Moteurzine »

[Désinscription](#) de  
« Moteurzine »