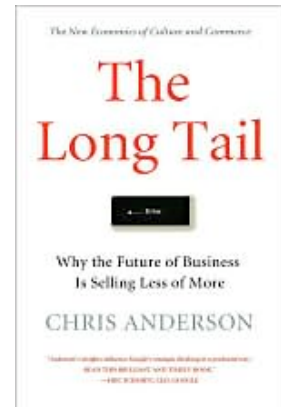


Longue Traîne et Référencement : un concept aujourd'hui incontournable

[Retour au sommaire de la lettre](#)

Le concept de Longue Traîne a été mis en lumière en 2004 par Chris Anderson, rédacteur en chef du site Wired, et tend à prouver que les articles les plus populaires, sur un site de vente en ligne, ne génèrent qu'une partie minoritaire du chiffre d'affaires, l'immense majorité de celui-ci étant pourvu par les très nombreux articles achetés uniquement à quelques exemplaires, l'effet de masse jouant à plein. Nous avons pu nous apercevoir que cette théorie s'appliquait également parfaitement au référencement et au trafic généré par les moteurs de recherche, à tel point qu'il semble indispensable aujourd'hui de prendre en compte cette notion de Longue Traîne dans une stratégie de visibilité sur les moteurs de recherche. Explications...

On doit la notion de "longue traîne" (ou *long tail* en anglais) à Chris Anderson, rédacteur en chef de la revue Wired, qui a développé ce concept dans un premier article pour la revue dont il s'occupe, puis sur son site web (<http://www.thelongtail.com/>) ainsi que dans un livre intitulé "The Long Tail" (traduit en français également mais la plupart des personnes qui ont lu les deux versions semblent avoir préféré l'original).



Principe de la longue traîne

Le principe de la longue traîne est apparu à Chris Anderson lorsqu'il a pu explorer les statistiques de ventes de sites web de commerce électronique comme Rhapsody (téléchargement de musique), Netflix (vente de DVD) ou Amazon. En inspectant une courbe présentant en abscisse les produits vendus et en ordonnée le nombre de ventes, il s'est rapidement aperçu que la courbe pour chacun des sites étudiés ressemblait à ceci :

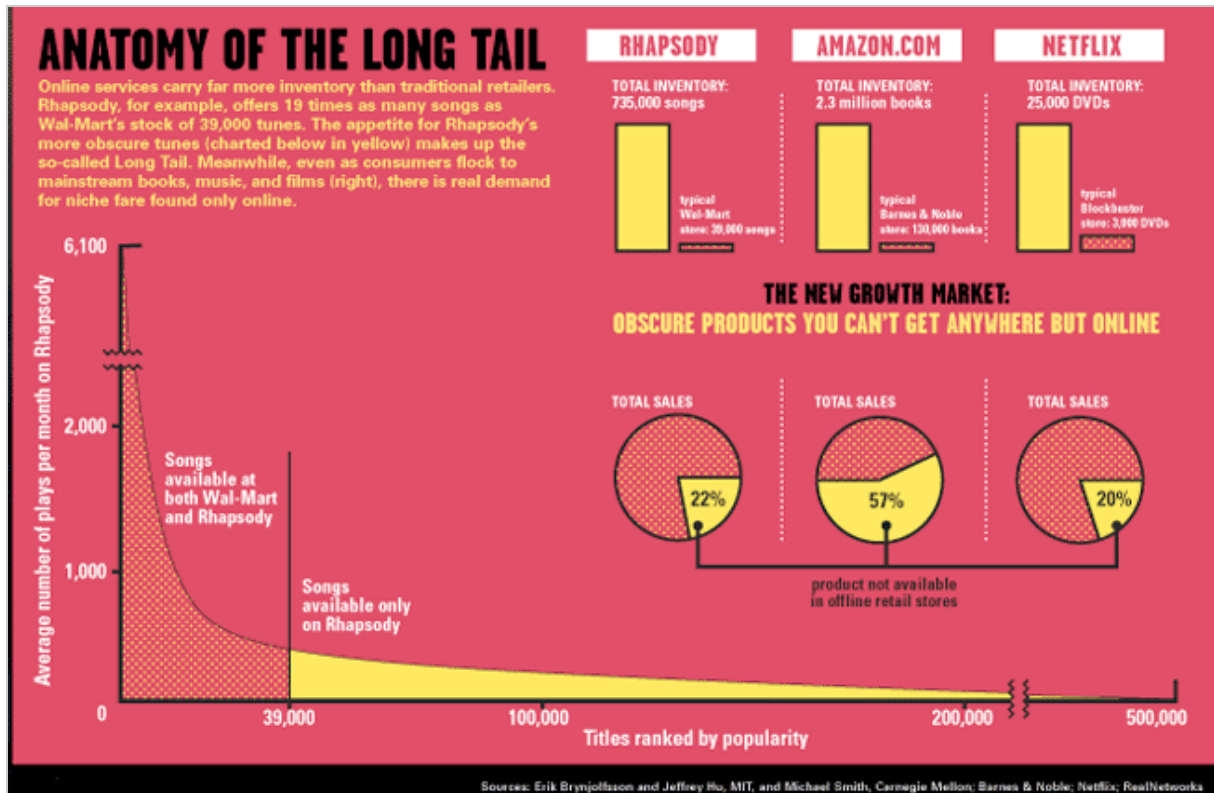


La partie rouge (la "tête" de la Longue Traîne) représente les best-sellers : peu de produits très populaires représentant de très nombreuses ventes. La partie jaune ("queue" de la Longue Traîne) est représentative de produits peu vendus (parfois une à deux ventes par mois) mais en très grand nombre.

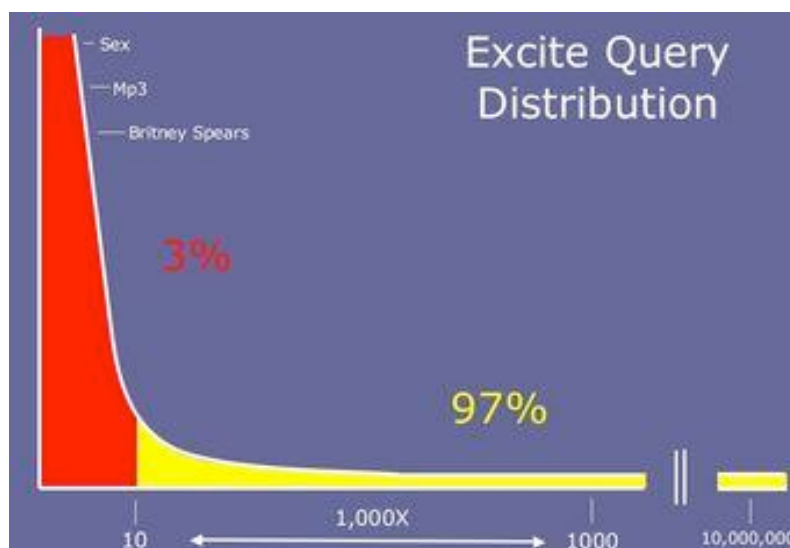
Le concept de Longue Traîne devient très intéressant lorsqu'on s'aperçoit qu'en fait, c'est la "queue" de cette Longue Traîne qui génère le plus souvent 80% du chiffre d'affaires de ce type de

site... C'est ainsi un très grand nombre de produits vendus très peu souvent qui, par leur masse, représente la majeure partie du bénéfice d'un site web tels que ceux inspectés par Chris Anderson.

De plus, si ce concept se vérifie pour, par exemple, des livres papier, il n'en reste pas moins vrai qu'il existe un besoin impératif de stockage de ces exemplaires papier, et ce stockage représente un coût... L'idée de Longue Traîne est donc encore plus intéressante pour des produits numériques (films, musique, études au format PDF, etc.) pour lesquels le coût de stockage est aujourd'hui quasi nul... Si vous arrivez à proposer en ligne tout ce qu'il est possible d'écouter en termes de musiques sur la planète, soit des millions, voire des milliards de morceaux, le constat de la Longue Traîne fait que vous pouvez être quasiment sûr que chacun d'entre eux sera au moins acheté une fois sur une année, générant ainsi un chiffre d'affaire considérable...



Autre exemple : Chris Anderson, dans son livre, explore également rapidement le monde des moteurs de recherche, notamment en regardant les statistiques des mots clés saisis sur le moteur Excite en 2001.



Il s'est rapidement aperçu que, chaque mois, 3% de toutes les recherches effectuées sur ce moteur se focalisaient sur une dizaine de mots ("sex", "mp3", "britney spears", etc.), le reste des requêtes était réparti sur des dizaines de millions d'autres termes et expressions... Les moteurs de recherche sont donc typiquement des outils basés sur un concept de "longue traîne". D'ailleurs, le marché publicitaire sur les liens sponsorisés, permettant de mettre en place des enchères potentiellement sur des millions de mots, est là aussi un marché-type de Longue Traîne... Et, quelque part, les moteurs de recherche, basés sur l'exploration des pages web, ne sont-ils pas la Longue Traîne des annuaires, qui n'inspectent que des fiches descriptives de sites ? Eric Schmidt, PDG de google, n'a-t-il pas dit, lors de la première assemblée générale des actionnaires du moteur, que la mission de sa société était d'être "au service de la Longue Traîne" ? On ne saurait être plus clair...

La Longue Traîne et le site Abondance.com

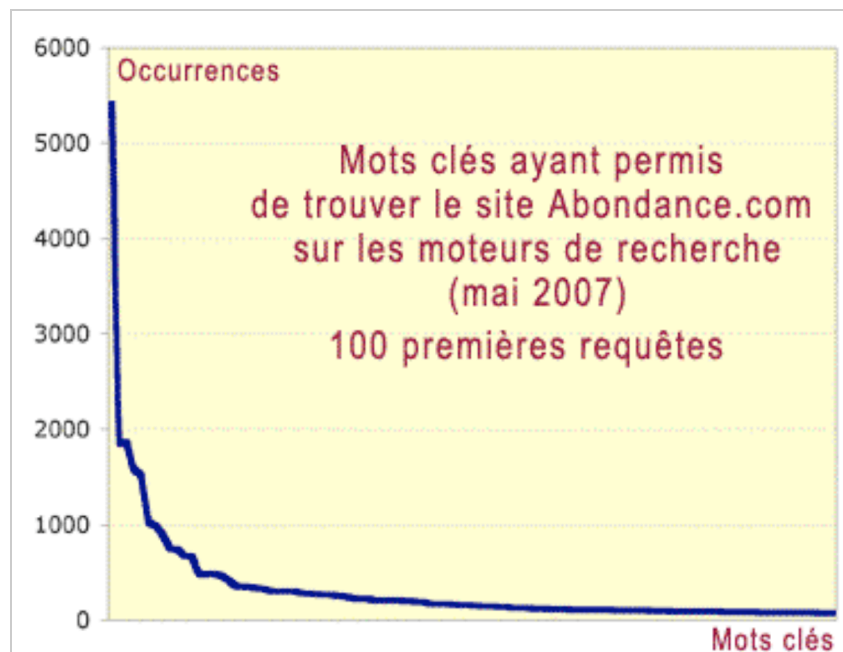
Nous venons d'explicitier rapidement (nous ne pouvons que vous encourager à lire le passionnant livre de Chris Anderson pour en savoir plus) ce concept de Longue Traîne qui révolutionne actuellement plusieurs concepts de l'économie numérique... Mais il existe un point qui n'est pas abordé (à moins que nous ne l'ayions manqué lors de notre lecture de l'ouvrage...), c'est l'application de ce concept au monde du référencement.

Ainsi, nous avons dans un premier temps inspecté les statistiques "mots clés" du site "abondance.com". Dans l'immense majorité des outils statistiques disponibles sur le marché, il existe une catégorie d'informations indiquant avec quels mots clés les internautes ont trouvé votre site sur les moteurs de recherche. C'est le cas de notre outil de stats (même s'il était limité par défaut aux 1 000 mots clés les plus populaires, nous avons pu retrouver les stats exhaustives que le système avait bien sauvegardé dans ses fichiers de données). Nous avons ainsi recoupé des informations sur les **25 296 requêtes** qui avaient permis de trouver le site web en question sur les différents moteurs de recherche du Web pendant le mois de mai 2007.

Il s'est rapidement avéré que :

- Moins de 10 requêtes ont généré chacune plus de 1% du trafic "moteurs" : *abondance* (6,4%), *miserable failure* (2,2%), *virtual earth* (2,2%), *moteurs de recherche* (1,8%), *robots.txt* (1,7%), *google video* (1,2%), *youtube* (1,1%), *humour* (1%).

- Si l'on ne prend en compte que les 100 premières requêtes les plus populaires, la courbe est déjà typiquement celle d'une Longue Traîne :



- Si l'on prend, cette fois, les 1 000 premières requêtes, le phénomène est toujours le même :



- Les 10 premiers mots clés ont généré 19,77% du trafic "moteur" (soit 16 637 visites sur les 84 124 amenées par les moteurs de recherche sur le mois testé), les 80,23% restant étant constitué par 25 286 expressions différentes, représentant chacune moins de 0,8% du trafic global... Un pur phénomène de "Longue Traîne"...

- La dernière requête à générer au moins 0,1% du trafic (*moteur recherche video*) occupe le 84ème rang. Ainsi, à partir de la 85ème expression, elles génèrent toutes moins de 0,1% du trafic "moteur" total...

- La 1 000ème requête identifiée (exemples, du 990ème au 1 000ème rang : *solution paiement en ligne, recherche en langue arabe, msn recherche, mon google, barre d'outils, localisé, moteurs de recherche vidéos, référencement sur moteur de recherche, annuaire besançon, le top, top mot clé*) génère encore 7 visites par mois sur le site...

- La 4 000ème requête (*adsense rss*) est encore demandée 2 fois dans le mois...

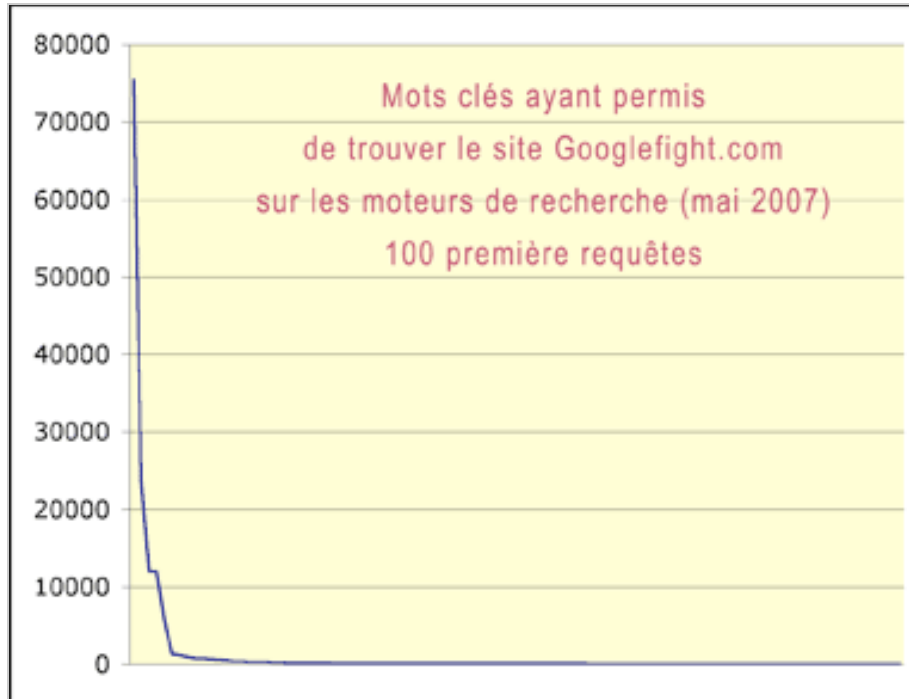
Voici le nombre de fois où le site a été trouvé pour des mots clés sur les moteurs de recherche :

Nombre de fois où le mot clé est demandé	Nombre de mots clés et expressions
1 occurrence	19 334
2 occurrences	2 931
3 occurrences	927
4 occurrences	499
5 occurrences	304
...	...

Ainsi, le site a été trouvé, sur un mois, 19 334 fois grâce à un mot clé qui a généré une visite unique. Une paille qu'il est difficile de négliger en termes de trafic...

La Longue Traîne et le site Googlefight.com

Nous avons effectué le même travail sur un autre site du Réseau Abondance, Googlefight.com, un jeu autour de Google, plus orienté "Grand public" et qui génère bien plus de visites (6 millions de pages vues par mois en moyenne) qu'Abondance qui est pour sa part plutôt voué à un public professionnel... De plus, ce site contient très peu de contenu, contrairement à Abondance.com. Le phénomène observé est pourtant exactement le même et les courbes obtenues strictement identiques à celles typiques de la Longue Traîne :



A ceci près que les requêtes "googlefight" et "google fight" représentent cette fois près de 70% des requêtes (on comprend ici l'influence sur ces chiffres du manque de contenu du site)... En revanche, les chiffres descendent très vite ensuite avec la 200ème recherche qui n'est plus demandée que 8 fois dans le mois, puis pour des milliers d'expressions qui ne génèrent qu'une ou deux visites mensuelles... En revanche, le nombre total de requêtes ayant généré au moins une visite sur le site *via* les moteurs de recherche (quelques milliers) est très éloigné des 25 296 du site Abondance.com. Absence de contenu oblige...

Ainsi, en étudiant les statistiques des sites Abondance.com et Googlefight.com, nous avons pu nous apercevoir à quel point la théorie de la Longue Traîne pouvait très facilement s'appliquer au référencement et aux statistiques de trafic sur son site généré par les moteurs de recherche. D'autres recherches et études que nous sommes en train d'effectuer dans un cadre plus large (qui devrait être dévoilé prochainement) sur d'autres sites web, bien plus connus qu'Abondance, semblent aujourd'hui donner exactement les mêmes résultats statistiques. Une bonne raison pour tenter d'extrapoler ce phénomène dans le cadre d'une stratégie de référencement globale...

Extrapolation de la Longue Traîne dans le cadre d'une stratégie de référencement

Lorsqu'on met en place une stratégie de référencement pour un site web, l'un des premiers réflexes que l'on a le plus souvent est de définir les mots clés sur lesquels on va tenter de positionner son site. En faisant cela, on va en fait "nourrir" la tête de la Longue Traîne avec des mots clés que l'on désire avant tout voir comme des "best-sellers". C'est, en revanche, le contenu textuel de votre site qui va nourrir la queue de la Longue Traîne...

Ainsi, si l'on désire optimiser son référencement en prenant en compte ce phénomène de "Longue Traîne", il faudra :

- Choisir ses mots clés de départ avec soin, sachant qu'il y a de fortes chances pour que ces mots clés choisis pour le positionnement ne représentent que 20% du trafic généré par les moteurs de recherche. Cependant, il y a également de fortes chances pour que ce trafic soit bien ciblé et de

"bonne qualité" puisqu'il correspondra à des mots correspondant à votre activité et que vous avez choisis comme étant pertinents pour trouver un site comme le vôtre... A vous donc d'optimiser certaines pages de votre site pour les mettre en valeur. La stratégie à adopter est donc, chronologiquement parlant :

- * Choix des mots clés ;
- * Optimisation de certaines pages du site en fonction de ces mots clés ;
- * Mesure du trafic généré spécifiquement sur ces requêtes.

On parlera ici de "trafic maîtrisé"...

- Optimiser la structure des pages de vos sites pour qu'elles mettent en valeur leur contenu éditorial afin de favoriser la "queue" de la Longue Traîne. Dans ce cas, vous ne maîtrisez, par définition, pas vraiment les positionnements obtenus, donc un certain pourcentage de ces expressions générera un trafic "stérile". Mais il y a fort à parier que, dans le "lot", certains mots clés seront très intéressants, même s'ils n'ont pas été imaginés ainsi dès le départ. Dans ce cadre, la stratégie à mettre en œuvre est différente :

* Optimisation du code HTML (titre en balises H1, adéquation entre balise Title, contenu de la balise meta Description, titre éditorial, texte de la page, importance du premier paragraphe de texte, etc.) des "masques" ou "templates" de vos pages selon les règles du "3C" (voir article à ce sujet dans cette lettre).

* Ecriture du contenu si possible en tenant compte des contraintes des moteurs (titres descriptifs, chapos contenant les mots clés importants pour bien comprendre le contenu de la page, mots clés importants en gras, intitulés de liens explicites, etc.).

* Mesure du trafic global généré par ces termes "secondaires"...

Le trafic dans ce cas devient "opportuniste" (ce qui n'a rien de péjoratif...)...



On le voit, les deux stratégies sont très complémentaires...

Il est en tout cas certain que le concept de Longue Traîne ne peut pas, en 2007, être ignoré dans le cadre d'un référencement. Il faut absolument, pour mettre en place une stratégie gagnante de visibilité sur les moteurs de recherche, soigner à la fois les mots clés qui vont nourrir la **tête** comme la structure du site qui va mettre en avant les mots clés de la **queue** de cette "**Long Tail**". Un travail certes parfois long mais qui peut rapidement porter ses fruits de façon très efficace...

Réagissez à cet article sur le blog des abonnés d'Abondance :
<http://abonnes.abondance.com/blogpro/2007/09/longue-trane-et-rfrencement-un-concept.html>