

*Etude sur le référencement des
livres et des auteurs sur moteur
de recherche en 2012*
2^e édition

E-MEDIATION

LE

RÉFÉRENCEMENT

DES LIVRES SUR

INTERNET n° 2

Septembre 2012

Auteur
Mathias Daval
(Edysseus Consulting),
pour le MOTif

le MOTif
Cécile Moscovitz
Responsable des études
6, villa Marcel-Lods
Passage de l'Atlas
Paris 19^e
Métro Belleville
01 53 38 60 61
contact@lemotif.fr

www.lemotif.fr
le MOTif
est un organisme associé
de la Région Ile-de-France.

 **île de France**

Introduction

La présente étude fait suite à notre première étude « Le Référencement des livres sur les moteurs de recherche » parue en octobre 2010¹ : elle propose à la fois une mise à jour des données (juin 2012) ainsi qu'un approfondissement de la recherche. Rappelons que, dans le cadre de nos études sur le livre numérique (programme « ElabZ »), l'objectif est de décrire les nouveaux leviers d'intermédiation, entre les consommateurs et les livres, spécifiques à l'écosystème numérique.

Pourquoi étudier le référencement ?

Nous avons analysé dans quelle mesure cette « e-médiation », qui passe par le référencement dit « naturel » (par opposition au référencement payant)² sur moteur de recherche, donne à l'internaute des informations ainsi que la possibilité de télécharger ou d'acheter un livre, qu'il soit papier ou numérique. Le but du présent tableau de bord est donc de cerner la capacité des différents acteurs de la filière du livre à obtenir un référencement efficace au sein des moteurs de recherche. Les questions auxquelles nous souhaitons répondre sont les suivantes :

- Quels sont les sites les mieux référencés ?
- Quelle est la diversité des sites proposés lorsqu'un internaute cherche un livre ou un auteur ?
- Le référencement est-il variable selon les mots-clés utilisés ?
- Quelle est la présence de sites « pirates » proposant une offre illégale de livres et parasitant ainsi les premiers résultats ?
- Quelle est la présence des acteurs traditionnels du livre ?

Si la méthodologie générale de « Référencement 1 » a été conservée, on peut noter que :

- Compte tenu de sa très forte domination dans les usages de recherche et des faibles différences, identifiées dans « Référencement 1 », entre Google, Yahoo! et Bing quant aux résultats de recherche générés, seul Google a été utilisé ici.
- Les 10 premiers résultats, extraits en mai et juin 2012, ont été analysés pour chaque requête (contre les 3 premiers dans « Référencement 1 »), soit la 1^{re} page de résultats.
- Le panel de livres issus des tops des ventes de *Livres-Hebdo* (« Corpus n° 1 », cf. tableau n° 20, p. 20) a été comparé à un échantillon de livres au succès commercial plus hétérogène (« Corpus n° 2 », cf. tableau n° 21, p. 21).
- Une série de mots-clés a été utilisée pour affiner la modélisation de la recherche faite en vue d'un achat.

Pour une introduction générale sur le marché et les usages du référencement en France, nous invitons le lecteur à se reporter à la première partie de « Référencement 1 » et à consulter notamment les études suivantes :

- *Référencer les œuvres sur Internet, prescrire la culture ?* (Forum d'Avignon, nov. 2011)³ ;
- Baromètre des moteurs de recherche (AT Internet)⁴ ;
- Etude sur le référencement de Wikipedia dans Google (Intelligent Positioning)⁵.

¹ Que nous désignerons dorénavant par « Référencement 1 ».

² Le référencement naturel est le positionnement d'un site sur les moteurs de recherche hors promotion payante, qui peut être amélioré par un ensemble de techniques dites de « SEO » (Search Engine Optimization ; voir aussi http://fr.wikipedia.org/wiki/Optimisation_pour_les_moteurs_de_recherche).

³ http://www.forum-avignon.org/sites/default/files/editeur/Brochure_BNP_ATELIER_2012_FR.pdf

⁴ <http://www.atinternet.fr/documents/barometre-des-moteurs-de-recherche-juin-2012/>

⁵ <http://www.intelligentpositioning.com/blog/2012/02/wikipedia-page-one-of-google-uk-for-99-of-searches/>

Conclusion et synthèse

- La recherche des livres et des auteurs sur les moteurs de recherche sur Internet révèle à la fois une **diversité des sites référencés**, ainsi qu'une **concentration** des sites : les **10 sites les plus référencés représentent 37 % des liens** proposés par Google sur l'ensemble des requêtes.
- Globalement, **les sites d'e-commerce représentent 40% des liens référencés**. Les sites d'e-commerce historiques **Amazon.fr et Fnac.com** occupent toujours une position dominante dans le référencement (**22,1 % des liens référencés**), quels que soient les mots-clés utilisés dans la recherche. **Decitre.fr** fait une entrée dans le top 5 des sites les plus référencés, et **Starzik** obtient un bon référencement avec une présence récurrente dans le top 10 sur plusieurs mots-clés. **Dialogues** fait son apparition dans le top 10 sur une recherche avec le mot-clé « ebook ». Enfin les distributeurs/librairies numériques comme **Epagine** et **Immateriel** figurent également dans le classement sur des recherches avec les mots-clés « ebook » et « pdf ».
- Les **sites d'informations édités par des tiers, en particulier les blogs de lecteurs**, forment une catégorie particulièrement présente (**21 % des sites référencés**). Les sites communautaires et les réseaux sociaux, en particulier **Babelio**, ont fait une percée remarquable dans le classement.
- **Les éditeurs et les auteurs sont correctement référencés via leurs sites officiels** (en cumulant les catégories « Editeurs » et « Auteurs », ils sont au 5^e rang des types de sites référencés) et peuvent ainsi être au contact des internautes lors d'une recherche d'un livre. Toutefois, des efforts doivent encore être faits dans un but d'optimisation du référencement ; ainsi ils apparaissent rarement dans les 3 premiers résultats, qui concentrent pourtant l'essentiel de l'attention de l'internaute.
- La **question du piratage** reste ambivalente : d'une part, à partir d'une recherche simple sur le titre d'un livre ou le nom d'un auteur, les sites pirates sont mal référencés par Google ; mais d'autre part, ces derniers sont assez présents sur des recherches plus spécifiques avec des mots-clés comme « ebook », « télécharger » et « pdf » (10 % des sites référencés).
- **Librairies et bibliothèques** sont peu référencées (moins de 1 % des résultats), pour des raisons à la fois structurelles liées au fait que leur site n'est pas marchand (l'optimisation du référencement n'est alors pas toujours leur objectif) et sans doute à la faiblesse des moyens techniques dont elles disposent.

Quelques pistes pour les acteurs de la chaîne du livre

De ces résultats, il est possible pour l'ensemble des acteurs de la filière du livre de définir des stratégies de présence et de promotion numérique afin d'être au plus près des internautes lors de la recherche d'un ouvrage.

- Pour les éditeurs et les auteurs, **l'utilisation des réseaux sociaux et des relais promotionnels via les blogs, les plates-formes communautaires** et les annuaires spécialisés est un passage incontournable pour amener le lecteur vers l'ouvrage qu'il recherche grâce au référencement naturel sur les moteurs de recherche. Si tel est le but recherché, il apparaît ainsi nécessaire de démultiplier les efforts de promotion au-delà des relais médias traditionnels afin de s'étendre à l'ensemble des sites d'infos dédiés au

livre au sein de l'écosystème numérique. En guise de réserve par rapport au « diktat » des moteurs de recherche, il convient de souligner que, dans le cadre de la **recommandation**, notamment lorsqu'elle passe par les réseaux sociaux, l'e-médiation court-circuite souvent l'usage des moteurs de recherche ;

- Les **contenus multimédias** sont valorisés par Google car référencés sous la forme de « push » vidéo ou image ;
- Amélioration des **métadonnées** (sur la fiche du livre, et au sein même de la page HTML), afin d'optimiser l'affichage des informations sur les pages de résultats de Google, et d'associer correctement chaque page de contenu sur un livre à un titre, un auteur et un éditeur ;
- Le **référencement payant** (AdWords) semble encore peu utilisé par les éditeurs comme les distributeurs, la publicité ne faisant globalement pas partie des habitudes de l'édition. La question de l'utilisation de la promotion payante en ligne, dans le cadre de campagne ciblées (AdWords ou sur les réseaux sociaux, notamment Facebook) est un des enjeux sur lesquels les éditeurs doivent s'interroger ;
- En ce qui concerne les **librairies et les bibliothèques**, la création d'un portail commun (de vente et/ou d'information) permettant de mutualiser les efforts et d'améliorer la visibilité reste une question d'actualité.

Méthodologie

Panel de livres

Les recherches ont été effectuées à partir des requêtes sur Google du panel des **75 livres** suivants (la liste complète se trouve en annexe page 20):

- **Corpus n° 1 : tops 10 des ventes de 2011** (*Livres-Hebdo*) sur 5 grandes catégories de livres, soit 50 titres : Romans, Essais, Jeunesse, Pratique, BD. Nous avons, lorsque cela a été possible, sélectionné les 10 premiers titres disponibles en version numérique légale (au total : 39 titres sur 50).
- **Corpus n° 2 : 25 titres provenant d'un panachage** de divers tops de ventes (top Datalib⁶, tops des ventes sur les plates-formes Eyrolles.com et Didactibook⁷ à la mi-juin 2012), ainsi que d'une sélection de titres représentatifs de la diversité éditoriale et disponibles en numérique.

Ce panel permet donc d'offrir un échantillon de « best-sellers » complété par des titres plus divers. Il n'est pas représentatif de l'ensemble des livres disponibles mais fournit le moyen de tester l'accessibilité de titres et d'auteurs moins populaires que les best-sellers de l'année. Cela permet de compenser le fait que notre corpus met l'accent particulièrement sur les plus gros acteurs en termes de distribution numérique, qui possèdent les moyens d'un référencement optimisé.

Modélisation du comportement de recherche des internautes

Afin de pouvoir étudier les résultats fournis par les moteurs de recherche, il est nécessaire de modéliser le comportement des internautes. Nous avons conservé la distinction entre 2 grands types de recherche :

- une recherche sur le **titre du livre**
- une recherche sur le **nom de l'auteur**

Pour chacun de ces types de recherche nous avons affiné la modélisation en ajoutant une série de mots-clés, parmi les plus communément utilisés lors des recherches de livres dans le but d'effectuer un achat ou un téléchargement : « ebook », « gratuit », « télécharger », « acheter » et « PDF ».

Ainsi, nous avons différencié 12 protocoles de recherche différents, correspondant à l'ajout après le titre ou le nom de l'auteur d'un mot-clé fréquemment utilisé par les internautes et tel que nous avons pu le vérifier d'après Google Trends et Google Suggest. Par exemple, pour le titre *L'appel de l'ange*, de Guillaume Musso, nous avons recherché sur Google.fr :

- | | |
|--|--------------------------------------|
| ▪ <i>L'appel de l'ange</i> | ▪ <i>Guillaume Musso</i> |
| ▪ <i>L'appel de l'ange ebook</i> | ▪ <i>Guillaume Musso ebook</i> |
| ▪ <i>L'appel de l'ange gratuit</i> | ▪ <i>Guillaume Musso gratuit</i> |
| ▪ <i>L'appel de l'ange télécharger</i> | ▪ <i>Guillaume Musso télécharger</i> |
| ▪ <i>L'appel de l'ange acheter</i> | ▪ <i>Guillaume Musso acheter</i> |
| ▪ <i>L'appel de l'ange PDF</i> | ▪ <i>Guillaume Musso PDF</i> |

⁶ Ventes pour la semaine du 4 au 8 juin 2012, d'après un panel d'environ 200 librairies indépendantes de premier niveau.

⁷ Titres disponibles en numérique sur : http://www.eyrolles.com/Accueil/Meilleures_ventes/ et http://www.didactibook.com/meilleures_ventes

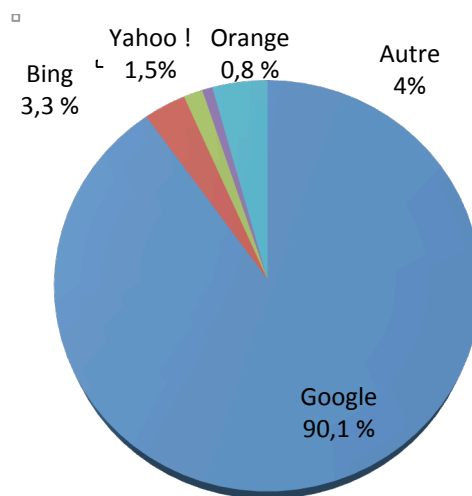
Compte tenu des constatations fournies dans « Référencement 1 » (la très faible distinction des résultats), nous avons abandonné l'usage d'un niveau d'expertise simple vs. un niveau avancé, correspondant à l'utilisation de guillemets (ex. : « *L'appel de l'ange* »). Toutes les requêtes ont donc été effectuées sans guillemets.

Cette modélisation permet de couvrir la majorité des modalités de recherche d'un internaute et d'étudier la stabilité des résultats en fonction de son type de recherche.

Choix des moteurs de recherche

Nous limitons notre extraction de données au principal moteur de recherche en France, Google, qui représente toujours **90 %** des visites en France.

Part de marché des moteurs de recherche en France en part de visites



Source : AT-Internet Institute (décembre 2011)

Extraction des résultats des moteurs de recherche

Pour l'ensemble des livres de notre panel et en fonction de l'ensemble des requêtes modélisées⁸, nous avons extrait (juin 2012) grâce à un script spécifique les 10 premiers liens fournis par Google, c'est-à-dire la première page de résultats.

Entre « Référencement 1 » et la présente étude, nous avons en effet décidé de faire passer la sélection à 10 résultats, afin d'étudier plus en détail le référencement de chaque catégorie de sites, y compris ceux qui n'apparaissent jamais dans les 3 premiers résultats. Toutefois, il convient de souligner que, selon les études effectuées par Google lui-même, seul 1 internaute sur 3 va au-delà des trois premiers résultats de recherche, et que le 1^{er} résultat est 4 fois plus cliqué que le 2^e.

⁸ En éliminant les requêtes par nom d'auteur en doublon de notre panel de livres. Par exemple « Anne Robillard » est présente deux fois dans notre échantillon.

Qualification des résultats de recherche

Etape 1 – Qualification automatique

Pour chacun des résultats (7 913), 2 critères ont été automatiquement attribués :

- leur visibilité, correspondant au rang du lien de 1 à 10
- l'adresse racine du site sur lequel ils sont hébergés
Ex. : http://fr.wikipedia.org/wiki/Guillaume_Musso est qualifiée « fr.wikipedia.org ».

Etape 2 – Qualification manuelle

Afin de pouvoir étudier de façon plus précise la catégorie des sites les mieux référencés, nous avons qualifié manuellement l'ensemble des liens extraits par notre script, et les avons regroupés en **20 catégories** :

Catégorie	Description
Wikipedia	Wikipedia et les sites relatifs (ex. : wikiquote).
Site Auteur ou Livre	Site officiel de l'auteur ou d'un livre (ex. : guillaumemusso.com).
Site Editeur	Site officiel de l'éditeur (ex. : albin-michel.fr).
E-commerce	Plate-forme d'e-commerce, qu'il s'agisse d'un libraire ou d'une GSS (ex. : Fnac.com, Amazon.fr, Epagine.fr).
Librairie	Site d'une librairie physique, qui ne possède pas d'activité d'e-commerce.
Infos Livres	Site d'informations dont le principal contenu éditorial concerne les livres (ex. : actualitte.com).
Infos BD	Site d'informations dont le principal contenu éditorial concerne les BD (ex. : bdgest.com).
News	Site d'actualités et d'infos générales (presse, radio, TV...). (ex. : tf1.fr)
Blog Personnel/Site de fan	Un blog ou un site, qu'il soit entièrement dédié ou non à un livre et/ou un auteur, réalisé par un lecteur ou un groupe de lecteurs.
Push Google	Proposition de contenus multimédias par Google (en général provenant de Google Images).
Site Avis	Site d'avis d'internautes (ex. : babelio.com).
Réseaux Sociaux	Réseau social généraliste (ex. : facebook.com).
Contenu Multimédia	Site agrégateur de photos, vidéos (ex. : youtube.com).
Infos Ciné	Site d'informations sur le cinéma (ex. : allocine.fr).
Bibliothèque	Site d'une bibliothèque ou d'une médiathèque.
Site Pirate	Site qui fournit l'accès à des contenus pirates liés à la requête.
Comparateur Prix	Comparateur de prix (ex. : ciao.fr).
Annuaire/Agrégateurs	Annuaire de ressources (ex. : digg.com).
Non Pertinent	Site non pertinent au regard de la requête réalisée (ex. : un site sur la série animée <i>Naruto</i> et qui ne parle pas de la BD).
Autre	Ce qui n'entre dans aucune des catégories ci-dessus.

Résultats

Les résultats doivent être considérés avec la précaution suivante : la fréquence d'apparition des différents sites est nécessairement liée à la structuration de notre échantillon de livres. Certaines url peuvent apparaître bien placées dans les classements de référencement tout simplement parce que l'auteur est présent plusieurs fois dans notre panel. Toutefois, on constate que ce biais reste marginal et ne modifie pas l'ordre des têtes de classement.

Diversité du référencement et concentration des sites

Nous avons commencé par analyser le nombre de sites distincts proposés par Google, qui s'élève à près de **1 600 sur un total d'environ 8 000 résultats**.

3 premiers résultats de recherche	10 premiers résultats de recherche
501	1 594

Toutefois l'analyse de ces sites, si elle en montre la **diversité** qui traduit une sorte de « longue traîne » dans le référencement des livres et des auteurs, montre aussi un **fort degré de concentration**. Les 15 sites le plus souvent présents parmi les 10 premiers résultats correspondent à plus de 40 % de ces résultats de recherche.

10 premiers sites	15 premiers sites
37 %	41,1 %

A l'inverse, 61 % des sites référencés ne le sont qu'1 seule fois (et 75 % que 2 fois).

Les tableaux suivants mettent en évidence les sites les plus présents sur les 10 premiers résultats de recherche sur Google.

Tableau n° 1 : Recherche multicritères sur les 10 premiers résultats de Google (tous types de recherche confondus)

		%	Catégorie
1	fnac.com	11,6 %	E-commerce
2	amazon.fr	10,5 %	E-commerce
3	decitre.fr	4,5 %	E-commerce
4	wikipedia.org	3,6 %	Wikipedia
5	ebooks-gratuit.com	1,4%	Site Pirate
6	starzik.com	1,2%	E-commerce
7	lexpress.fr	1,1%	News
8	babelio.com	1,1%	Site Avis
9	evene.fr	1,1%	Infos Livres
10	ebookbrowse.com	0,9 %	Annuaire / Site Pirate

Comment lire le tableau ? « La Fnac représente 11,6 % des occurrences de lien de notre échantillon (soit 920 liens) référencés par Google entre le rang 1 et le rang 10. »

Tableau n° 2 : Recherche multicritères sur les 3 premiers résultats de Google (tous types de recherche confondus)

		%	Catégorie
1	fnac.com	18,2 %	E-commerce
2	amazon.fr	17,3 %	E-commerce

3	wikipedia.org	5,1 %	Wikipedia
4	decitre.fr	2,8 %	E-commerce
5	ebooks-gratuit.com	2,4%	Site Pirate
6	didactibook.com	1,3%	E-commerce
7	starzik.com	1,1%	E-commerce
8	itunes.apple.com	0,9 %	E-commerce
9	priceminister.com	0,9 %	E-commerce / Comparateur
10	lexpress.fr	0,9 %	News

Fnac, Amazon et Wikipedia, comme dans « Référencement 1 », restent donc le trio de tête.

Remarque : Didactibook, Itunes et Priceminister apparaissent dans le top 3 et pas le top 10 car ils ne font pas partie des sites qui apparaissent à plusieurs rangs différents dans le top 10 (plusieurs résultats pour une même requête).

Stabilité des résultats de recherche au regard du type de requête

Dans le cadre de notre étude, nous avons modélisé le comportement de recherche des internautes selon les critères de recherche déjà évoqués.

Tableau n° 3 : Recherche « titre » sur les 10 premiers résultats de Google

	URL	%	Catégorie
1	amazon.fr	10,3 %	E-commerce
2	fnac.com	8,3 %	E-commerce
3	wikipedia.org	4,8 %	Wikipedia
4	babelio.com	2,5 %	Site Avis
5	lexpress.fr	2,1 %	News
6	push google	1,8 %	Push Google
7	decitre.fr	1,8 %	E-commerce
8	evene.fr	1,6 %	Infos Livres
9	booknode.com	1,4 %	Site Avis
10	allocine.fr	1,2 %	Infos Cinéma

Tableau n° 4 : Recherche « titre + télécharger » sur les 10 premiers résultats de Google

	URL	%	Catégorie
1	fnac.com	10,9 %	E-commerce
2	amazon.fr	6,7 %	E-commerce
3	decitre.fr	4,1 %	E-commerce
4	itunes.apple.com	2,9 %	E-commerce
5	wikipedia.org	2,9 %	Wikipedia
6	starzik.com	1,8 %	E-commerce
7	ebooks-gratuit.com	1,8 %	Site Pirate
8	bedetheque.com	1,7 %	Infos BD
9	t411.me	1,7 %	Site Pirate
10	fluket.com	1,5 %	Site Pirate

Tableau n° 5 : Recherche « titre + gratuit » sur les 10 premiers résultats de Google

	URL	%	Catégorie
1	amazon.fr	14,4 %	E-commerce
2	fnac.com	10 %	E-commerce
3	decitre.fr	4,8 %	E-commerce
4	evene.fr	2,2 %	Infos Livres
5	ebooks-gratuit.com	2 %	Site Pirate
6	facebook.com	2 %	Réseaux Sociaux
7	actuabd.com	1,2 %	Infos BD
8	blablaland.com	1,1 %	Autre

9	itunes.apple.com	1,1 %	E-commerce
10	lexpress.fr	1,1 %	News

Tableau n° 6 : Recherche « titre + acheter » sur les 10 premiers résultats de Google

	URL	%	Catégorie
1	amazon.fr	13,8 %	E-commerce
2	fnac.com	11,6 %	E-commerce
3	decitre.fr	3,4 %	E-commerce
4	priceminister.com	2,7 %	E-commerce / Comparateur
5	babelio.com	2,3 %	Site Avis
6	cdiscount.com	1,9 %	E-commerce
7	lexpress.fr	1,7 %	News
8	rueducommerce.fr	1,4 %	E-commerce / Comparateur
9	wikipedia.org	1,1 %	Wikipedia
10	evene.fr	1,1 %	Infos Livres

Tableau n° 7 : Recherche « titre + PDF » sur les 10 premiers résultats de Google

	URL	%	Catégorie
1	fnac.com	9 %	E-commerce
2	wikipedia.org	5,3 %	Wikipedia
3	banque-pdf.fr	3,5 %	Annuaire / Site Pirate
4	rechercher.me	2,7 %	Annuaire
5	decitre.fr	2,1 %	E-commerce
6	t411.me	2,1 %	Site Pirate
7	ebooks-gratuit.com	1,8 %	Site Pirate
8	ebookbrowse.com	1,8 %	Annuaire / Site Pirate
9	starzik.com	1,6 %	E-commerce
10	numilog.com	1,3 %	E-commerce

Tableau n° 8 : Recherche « titre + ebook » sur les 10 premiers résultats de Google

	URL	%	Catégorie
1	fnac.com	13,8 %	E-commerce
2	amazon.fr	9,7 %	E-commerce
3	decitre.fr	6 %	E-commerce
4	librairie.immateriel.fr	4,3 %	E-commerce
5	ebooks-gratuit.com	2,8 %	Site Pirate
6	ebookbrowse.com	2,2 %	Annuaire / Site pirate
7	starzik.com	2,1 %	E-commerce
8	lexpress.fr	1,9 %	News
9	numilog.com	1,8 %	E-commerce
10	librairiedialogues.fr	1,8 %	E-commerce

Tableau n° 9 : Recherche « auteur » sur les 10 premiers résultats de Google

	URL	%	Catégorie
1	wikipedia.org	8,5 %	Wikipedia
2	fnac.com	7,9 %	E-commerce
3	amazon.fr	6,1 %	E-commerce
4	babelio.com	4,1 %	Site Avis
5	decitre.fr	3 %	E-commerce
6	facebook.com	2,7 %	Réseaux sociaux
7	evene.fr	2,6 %	Infos Livres
8	rueducommerce.fr	2 %	E-commerce / Comparateur
9	lexpress.fr	2 %	News
10	push google	1,2 %	Push Google

A noter que ces classements sont identiques concernant le Corpus n° 2 seul. Cela signifie donc que les sites les mieux positionnés sur Google à partir d'une recherche de livre ou d'auteur (notamment les plates-formes d'e-commerce) conservent leur référencement qu'il s'agisse d'une recherche sur des best-sellers ou non.

Par ailleurs, les résultats des requêtes supplémentaires « auteur » + « mot-clé » sont sensiblement les mêmes que pour « auteur » seul, avec des variations déjà observées pour une recherche par « titre » : apparition de l'annuaire pirate Ebookbrowse pour une recherche avec « ebook » ou « PDF », bon positionnement du distributeur numérique Numilog sur une recherche avec « PDF », etc.

Principales observations

- **Un impact relatif des mots-clés** : au-delà du terme principal recherché (titre ou nom d'auteur), le ou les mot(s)-clé(s) ajouté(s) en complément ont globalement peu d'impact sur les premiers résultats de recherche. Toutefois, on constate qu'une recherche sur un auteur met davantage en avant Wikipedia (il y existe davantage de fiches sur les auteurs que de fiches sur les livres) ; et voir ci-dessous en ce qui concerne le piratage.
- **Le trio de tête des sites** les plus présents dans « Référencement 1 » (**wikipedia.fr, amazon.fr et fnac.com**) reste très bien positionné. Cela dénote la persistance de la domination de ces sites sur Google, et leur importance dans le processus d'information et d'achat des internautes dans le cadre d'une recherche sur les livres. A noter que **Decitre.fr** fait une remontée marquante dans le classement, qui témoigne de la qualité de son référencement récent (et traduit sa position dans les premiers cybervendeurs de livres en France, après Amazon et la Fnac).
- **Une très forte domination des sites d'e-commerce** : cf. analyse ci-après par catégorie éditoriale.
- **Bon positionnement des sites pirates**, présents dans la majorité des classements. La recherche par certains mots-clés, en particulier « PDF » et « télécharger », met en valeur la qualité de leur référencement et la concurrence qu'ils peuvent faire subir à l'offre légale. A noter que la moitié de ces sites sont des moteurs de documents au format texte (surtout pdf, doc et rtf), indexant à la fois du contenu pirate et du contenu légal.
- **Entrée de Facebook**, qui fait son apparition dans le top 10 de certaines requêtes, et montre l'importance croissante du référencement des réseaux sociaux dans la production de contenus sur les moteurs de recherche⁹. Globalement, Facebook se situe à la 12^e place des sites les mieux référencés (cf. annexes, tableau n° 22, p. 22), tous types de recherches confondus¹⁰. Les sites d'avis communautaires sont également bien positionnés, à l'image de **Babelio** qui se situe à la **4^e place** dans le classement général par titre (tableau n° 3, p. 7).
- **Ne pas négliger le « push Google »** : les contenus multimédias proposés par Google (essentiellement des images, plus marginalement de la vidéo ou de l'achat de livres) sont le **22^e** site/type de contenu le plus indexé dans le top 10 des recherches sur le total déjà

⁹ Deux types principaux de pages Facebook sont référencées dans nos résultats : les pages officielles d'auteurs (ex. : <http://www.facebook.com/guillaume.musso.fanpage>) et les pages d'auteurs ou d'œuvres générées à partir de leurs fiches Wikipedia (ex. : <http://www.facebook.com/pages/L%C3%89trange-Voyage-de-monsieur-Daldry/209059412448837>).

¹⁰ A noter que, au-delà des livres, sur toutes les recherches effectuées sur le web, Facebook génère à peine 0,6 % des références en septembre 2011 (selon NetMarketShare.com).

mentionné d'environ 1 600 sites différents (cf. annexes, tableau n° 22, p. 22), mais surtout apparaissent en bonne position sur certains types de recherche (4^e place pour les recherches par titres de roman, par exemple, cf. tableau ci-dessous). En termes de SEO, ce résultat constitue un enseignement pour l'ensemble des acteurs de la filière du livre. La promotion des sites ne se limite pas uniquement au déploiement de ressources éditoriales textuelles sur Internet mais concerne aussi la production de contenus multimédias spécifiques. Cette stratégie est d'autant plus efficace que, comme il a été mis en évidence dans « Référencement 1 », les contenus d'actualités et multimédias sont de plus en plus mis en avant par les algorithmes de recherche de Google et des autres moteurs de recherche.

Stabilité des résultats de recherche par segment éditorial de notre échantillon

Nous allons maintenant identifier les sites les mieux référencés au regard des segments éditoriaux de notre panel.

Tableau n° 10 : Recherche Romans

	Recherche par titre		Recherche par auteur	
1	amazon.fr	10,5 %	wikipedia.org	9,4 %
2	fnac.com	10,5 %	amazon.fr	8,8 %
3	babelio.com	8 %	babelio.com	8,1 %
4	push Google	6,2 %	fnac.com	6,9 %
5	evene.fr	5,6 %	evene.fr	6,3 %

Tableau n° 11 : Recherche Essais

	Recherche par titre		Recherche par auteur	
1	amazon.fr	10,9 %	wikipedia.org	8,3 %
2	fnac.com	6,7 %	fnac.com	5,1 %
3	wikipedia.org	4,2 %	amazon.fr	3,8 %
4	lexpress.fr	4,2 %	lexpress.fr	3,2 %
5	babelio.com	1,8 %	ruedecommerce.fr	3,2 %

Tableau n° 12 : Recherche Jeunesse

	Recherche par titre		Recherche par auteur	
1	wikipedia.fr	10,2 %	wikipedia.org	12 %
2	amazon.fr	8,8 %	fnac.com	8 %
3	booknode.com	6,8 %	babelio.com	5,6 %
4	fnac.com	6,8 %	zonelivre.fr	4,8 %
5	allocine.fr	3,4 %	amazon.fr	4 %

Tableau n° 13 : Recherche Pratique

	Recherche par titre		Recherche par auteur	
1	amazon.fr	10,8 %	decitre.fr	14,1 %
2	fnac.com	9,2 %	amazon.fr	12,9 %
3	decitre.fr	5,8 %	fnac.com	12,9 %

4	didactibook	2,5 %	facebook.com	4,7 %
5	laprocedure.com	2,5 %	marabout.com	3,5 %

Tableau n° 14 : Recherche BD

	Recherche par titre		Recherche par auteur	
1	amazon.fr	10,2 %	fnac.com	9,3 %
2	fnac.com	8,8 %	wikipedia.org	8,5 %
3	dargaud.com	5,8 %	bdencore.com	4,7 %
4	planetebd.com	3,6 %	actuabd.com	4,7 %
5	bdgest.com	3,6 %	bedetheque.com	3,9 %

L'analyse de la stabilité du référencement par segment éditorial de livres de notre panel montre que les sites les plus présents restent majoritairement identiques quels que soient le segment éditorial et le type de requête réalisée. On notera tout de même la possibilité pour quelques sites thématiques (notamment sur la BD) de s'insérer dans le palmarès des sites les plus présents sur les moteurs de recherche.

Analyse du référencement par catégorie de contenu

Tableau n° 15 : Classement des catégories par présence dans le référencement sur Google (tous types de recherche confondus)

	Catégorie de site	% top 10	% top 3
1	E-commerce	40	46,8
2	Site pirate	10,5	12,7
3	Blog perso	7,6	5
4	Non pertinent	7,3	4,7
5	News	6,1	2,8
6	Autre	5,1	3,1
7	Wikipedia	4	5,6
8	Site avis	3,5	2,4
9	Site éditeur	3,2	4,4
10	Push Google	3	3,2
11	Site auteur	2,4	4
12	Infos livres	2,4	1,3
13	Annuaire	2,1	1,7
14	Infos BD	1,6	0,7
15	Librairie	0,7	0,2
16	Réseaux sociaux	0,8	0,3
17	Contenu Multimédia	0,6	0,5
18	Bibliothèque	0,4	0,5
19	Comparateur prix	0,4	0,2
20	Infos ciné	0,2	0,2
	Total	100%	100%

Le tableau ci-dessus se lit ainsi : « Les sites d'e-commerce représentent 40 % des 10 premiers résultats sur Google pour l'ensemble des titres et auteurs de notre échantillon et 46,7 % des 3 premiers résultats. »

La catégorie « Non Pertinente » recouvre essentiellement les cas de figures suivants :

- Du contenu sans rapport avec l'œuvre (cas de titres composés de mots génériques) ;
- De l'information sur un homonyme (notamment en cas de nom d'auteur courant), en particulier sur les annuaires et les réseaux sociaux ;
- De l'information sur l'adaptation vidéo/ciné de l'œuvre sans aucun lien direct avec l'œuvre elle-même (fréquent pour les mangas, par exemple).

La catégorie « Autre » regroupe essentiellement des sites officiels d'organismes privés ou public (mairies, ministères, musées, entreprises...), notamment dans le cadre de newsletters consacrées aux publications récentes de livres, la venue d'auteurs pour des dédicaces, etc.

Tableau n° 16 : 10 catégories de site les mieux référencées lors d'une recherche « titre »

	Catégorie	% top 10	% top 3
1	E-commerce	26,1	30,6
2	Blog perso	14,1	9,3
3	News	11,1	6,8
4	Site éditeur	7,1	11,9
5	Push Google	6,6	7,6
6	Wikipedia	4,5	9,3
7	Infos livres	4,1	3,8
8	Site auteur	3,6	5,1
9	Infos BD	3,6	0,4
10	Librairie	0,9	0,7

Tableau n° 17 : 10 catégories de site les mieux référencées lors d'une recherche « auteur »

	Catégorie	% top 10	% top 3
1	E-commerce	21,1	19,1
2	News	13,9	5,2
3	Wikipedia	8,4	21,9
4	Blog Perso	7,9	5,2
5	Site auteur	7,5	18,6
6	Push Google	7,5	12,9
7	Infos livres	5,3	2,4
8	Site éditeur	4,7	5,2
9	Infos BD	3	1,4
10	Librairie	0,4	0,1

Observations

Globalement, les 10 catégories les plus référencées représentent 90,6 % du top 3 et 88,1 % du top 10.

Les plates-formes d'e-commerce dominant le référencement. La catégorie « e-commerce » reste la mieux référencée, avec le top 10 suivant :

Tableau n° 18 : top 10 des plates-formes d'e-commerce

1	Fnac	6	Numilog
2	Amazon (1)	7	Didactibook
3	Decitre	8	Immateriel
4	Starzik	9	Rueducommerce
5	Itunes	10	Cultura

(1) Cumul de Amazon.fr et Amazon.com

Amazon.fr et Fnac.com, à elles seules, représentent **22,1 %** des liens de référencement (cf. tableau n° 1). Ces plates-formes ainsi que les plus grandes plates-formes d'e-commerce présentes dans le top 10, comme **Decitre**, disposent en effet de plusieurs leviers d'optimisation de leur référencement sur Google :

- La taille de leur catalogue (plusieurs milliers de pages), qui génère des références quantitativement importantes dans l'indexation par Google ;
- La forte notoriété de leur enseigne auprès du public d'acheteurs de livres ;
- Des outils de web marketing et une expérience liés au référencement naturel. Les ressources investies en SEO (Search Engine Optimization) par les sites d'e-commerce historiques constituent une forte barrière à l'entrée pour tout nouvel entrant sur le marché de la distribution de livre physique ou numérique.

Cette catégorie inclut également d'autres acteurs dont la présence est plus faible, comme **Starzik** (qui assure un bon référencement, avec une présence récurrente dans le top 10 sur différents mots-clés recherchés), Dialogues (cf. top 10 du tableau 8), Gibert Jeune, Gibert Joseph, Le Furet du Nord, La Procure, ainsi que d'autres librairies indépendantes comme Brouillon de Culture (Caen) ou Cheminant (Vanne). Enfin la catégorie inclut les **distributeurs/librairies numériques** à l'instar d'Epagine ou d'Immateriel, dont la présence est marginale mais qui, par rapport à « Référencement 1 », font leur apparition dans les tops du référencement, en particulier lors de recherches avec certains mots-clés spécifiques (« ebook » et « pdf »).

On notera également la performance des **sites d'éditeurs** et des **sites officiels** (livre ou auteur) qui arrivent à se positionner lors d'une recherche simple ou avec mots-clés, qu'il s'agisse de titres ou de noms d'auteurs (5,6 % des résultats de recherche, cf. tableau n° 15, p. 11-12, soit au 5^e rang du référencement). Cela est encourageant pour les éditeurs, qui parviennent à être référencés dans les 10 premiers résultats malgré la présence écrasante des plates-formes d'e-commerce ; mais d'un autre côté, cela traduit aussi la nécessité d'améliorer leur action de SEO.

On remarquera ensuite que les **sites d'informations et de critiques** édités par des tiers (catégories « infos livres », « infos BD », « infos ciné », « blog perso », « site d'avis » et « news ») constituent une part conséquente du classement, avec **21 %** des sites les plus référencés (cf. tableau 15). Babelio.com et Evene.fr sont les deux sites qui tirent le mieux leur épingle du jeu avec un très bon positionnement sur la majorité des recherches, et un classement général respectivement à la 8^e et 9^e place des sites les plus référencés (cf. tableau 1). Dans la catégorie « News », c'est le site de *L'Express*, en 7^e position du classement, qui profite le mieux de son référencement naturel. Il est suivi, de plus loin, par *Télérama*, *Le Figaro* et *France Culture*. En ce qui concerne la BD, les sites d'informations spécifiques se retrouvent particulièrement bien référencés, en particulier les 5 premiers : Actua Bd, la Bédéthèque, Bd Gest, Bd Encre et Planète Bd.

A noter que la catégorie « **Blog perso** » recouvre des réalités différentes, depuis le mini-blog d'un lecteur qui publie sporadiquement des avis sur les livres qu'il lit, jusqu'à une communauté très organisée et très active réunie autour d'un forum consacré à un auteur ou une œuvre en particulier (ex. : largowinch-fr.com).

Si les **sites pirates** restent peu présents parmi les 3 et même les 10 premiers résultats de recherche lors de requêtes génériques par titre ou par nom d'auteur, l'ajout de mots-clés spécifiques comme « pdf », « télécharger » ou « gratuit » les fait remonter dans le classement. Le site **Ebooks-gratuit** fait également figure d'exception puisqu'il apparaît bien positionné dans presque tous les classements (recherche globale ou recherche avec chaque mot-clé). Globalement, sur l'ensemble des requêtes, les sites pirates se situent en 2^e place avec plus de 10 % des liens référencés (cf. tableau n° 15) ; toutefois, ce résultat n'est qu'une indication statistique qui traduit le fait que l'utilisation de certains mots-clés valorise la visibilité du piratage. A noter enfin que le téléchargement direct ou « direct download » y représente environ 70 % des liens contre 30 % pour le peer to peer.

Enfin, **Wikipedia** conserve une place de choix dans le référencement, particulièrement dans le cadre d'une recherche sur un nom d'auteur (avec près de 22 % des liens référencés sur les 3 premiers résultats de Google).

Les **librairies**¹¹ et les **bibliothèques** sont peu référencées, mais pas absentes du tableau (respectivement 0,7 % et 0,4 % des résultats), notamment les librairies lors des recherches par titre ou par auteur (rang n° 10). Leur « mauvais » classement traduit plusieurs réalités : le référencement de leurs catalogues sur internet n'est, pour la plupart d'entre elles, pas (ou pas encore) une priorité ; elles ne disposent pas des outils techniques et des moyens d'un référencement optimal ; leur grand nombre (contrairement au faible nombre de cybervendeurs, par exemple) parasite leur visibilité. Ces observations posent la question de la pertinence de l'existence d'un **portail commun** permettant de mutualiser les efforts et d'améliorer leur référencement, notamment après l'échec (en ce qui concerne les librairies) du site « 1001 Libraires ». La problématique des bibliothécaires comme « médiateurs numériques » est tout aussi d'actualité¹².

Si l'on prolonge l'analyse avec le **rang moyen** des principaux sites sur les 10 premiers résultats de Google (cf. tableau n° 1), on se rend compte qu'ils sont également les mieux référencés sur les 3 premiers résultats (à l'exception de Decitre.fr). Ainsi plus de 50 % des liens aboutissant sur Fnac.com, Amazon.fr ou Wikipedia.org sont positionnés parmi les 3 premiers rangs de résultats (cf. tableau n° 23 p.23).

L'observation spécifique du Corpus n° 2 apporte quelques remarques complémentaires.

Tableau n° 19 : Répartition des types de sites sur le Corpus n° 2

Catégorie	% top 10	% top 3
E-commerce	40,9	50,9
Blog perso	9,4	5,7
Non pertinent	8,9	4,7
News	7,1	4,1
Site éditeur	5,0	7,7
Infos livres	4,0	2,6
Site pirate	3,8	3,8
Site auteur	3,2	5,3
Autre	3,2	3,0
Wikipedia	3,0	3,6

¹¹ Rappelons qu'il s'agit de librairies physiques ne possédant pas de site d'e-commerce, par exemple, dans notre échantillon, la librairie Del Duca à Paris ou Le Préambule à Bayeux.

¹² <http://www.bibliobsession.net/2012/09/10/les-bibliothecaires-mediateurs-dans-locean-du-web/>. Notons que Gallica figure à la 45^e position des sites les mieux référencés (cf. tableau n°22 p.22).

- Faible taux de liens de téléchargement illégal (moins de 4 % contre 10 à 13 % pour le panel global), qui confirme l'observation faite par nos études « EbookZ » selon laquelle ce sont d'abord les best-sellers et les grosses maisons d'édition qui sont piratés ;
- Wikipedia un peu moins bien référencé : les éditeurs et auteurs moins connus ont moins de probabilité d'avoir leur fiche Wikipedia que les best-sellers.
- Meilleur positionnement de la catégorie « Blogs perso », qui confirme le fait que les relais d'information alternatifs sont très importants pour les auteurs et éditeurs ;
- Plus forte proportion de contenu « Non Pertinent », liée surtout à la structure de l'échantillon (et la présence de titres génériques dans le Corpus n°2).

Analyse du référencement par accessibilité de l'ouvrage

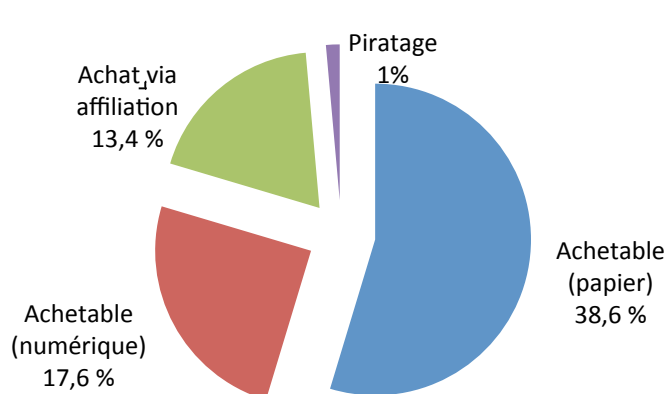
Nous nous sommes mis à la place d'un utilisateur cherchant à acheter ou télécharger un livre, numérique ou papier : il est donc important d'affiner l'analyse en observant si les liens fournis par Google, au-delà de la catégorie dans laquelle ils se situent, apportent ou non satisfaction à l'internaute.

Nous avons sélectionné 50 titres (30 titres provenant du Corpus n° 1, soit les 6 premiers de chaque catégorie, et 20 du Corpus n° 2), et pour chacun d'eux nous indiquons sur combien des 10 premiers liens proposés par Google (avec une recherche sur titre seul) le livre cherché est réellement achetable en version papier, numérique, ou téléchargeable gratuitement sur un site pirate. A noter que 44 titres sur les 50 (soit 88 % de cet échantillon) sont disponibles sur les plates-formes numériques légales.

49 titres sur 50 sont acheteables en 1 ou 2 clics grâce à au moins l'un des 10 premiers liens de résultats sur Google. Sur ces 49 titres, 32 peuvent être téléchargés légalement au format numérique.

Achetable	49 (dont numérique : 32)
Pas achetable	1

Part des liens référencés sur Google (top 10) permettant l'achat d'un livre



Près de 40 % des 10 premiers liens référencés par Google permettent l'achat du livre recherché sous forme papier. Un peu moins de 20 % des liens renvoient vers la possibilité d'un achat numérique.

En ce qui concerne l'affiliation, **13,4 %** des liens permettent de se procurer l'ouvrage de façon indirecte *via* un partenariat avec, en général, Amazon ou la Fnac (format papier ou numérique).

Enfin, on constate que le piratage est très peu représenté ici (et très inférieur à son taux de référencement global tel qu'il a été décrit plus haut). Cela s'explique de 2 façons :

- La recherche par « titre » seul le fait peut ressortir dans les 10 premiers résultats ;
- Les liens pirates référencés n'aboutissent que rarement à la possibilité réelle de télécharger l'œuvre (lien mort, recherche erronée, etc.).

Ces constats reflètent l'excellent positionnement des plates-formes d'e-commerce dans les moteurs de recherche, et la bonne accessibilité générale des ouvrages.

Observations propres au moteur de recherche de Google

Enfin, il convient de rappeler 3 critères propres à Google qui peuvent influencer l'accessibilité des ouvrages et des auteurs :

a) Les AdWords

Le programme AdWords permet aux annonceurs d'acheter de l'espace publicitaire sur Google associé à des mots-clés spécifiques, sous la forme d'annonces soit intégrées aux résultats sur un fond de couleur différencié, soit sous la forme de petits encarts latéraux. D'après l'analyse de notre échantillon, **40 %** des recherches par titre aboutissent à des publicités AdWords, provenant pour 75 % d'entre elles de Fnac.com, Amazon.fr et MyBoox.fr.

b) Le « push » Google Shopping

L'outil Google Shopping, de la même façon que Google Images, propose dans ses liens référencés un lien vers une page agrégeant des suggestions de contenus associés à la requête demandée. Dans le cas des livres, il s'agit de renvois vers des sites de libraires en ligne et d'e-commerce. Par exemple, une recherche sur le titre « Indignez-vous » fournit parmi les 10 premiers résultats un lien vers la page suivante :

The screenshot shows a Google Shopping search for the book 'Indignez-vous !'. On the left, there are filters for 'Actualités', 'Shopping', 'Livres', and 'Plus'. Below these are options to 'Afficher uniquement' (Livraison gratuite, Articles neufs), 'Toutes les catégories' (Livres), 'Tous les prix' (Jusqu'à 4 €, 4 € - 7 €, 7 € - 10 €, 10 € - 20 €, Plus de 20 €), and 'Tous les marchands' (Decitre, Gibert Joseph, Alapage, gibertjeune.fr, Amazon.fr, Plus). The main results list four items:

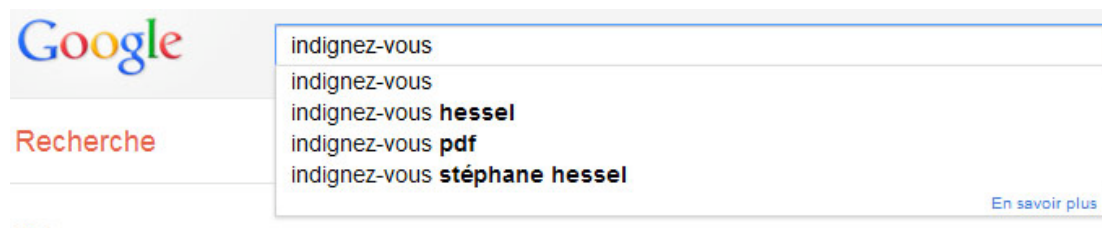
Product Title	Price	Additional Info
Indignez-vous ! édition revue et augmentée	2,94 €	Livraison gratuite Amazon.fr, 1 328 avis
Indignez-vous !	2,95 €	8,85 € (frais de port inc), La Procure
Indignez-vous !	2,95 €	Livraison gratuite Fnac.com
Indignez-vous !	2,94 €	4,94 € (frais de port inc), Mollat
Indignez-vous ! Edition revue et augmentée	2,95 €	Livraison gratuite Alapage, 38 avis

Seuls **12 %** des recherches par titre de notre échantillon aboutissent à une proposition de push shopping.

A noter que le Corpus n° 2 connaît des taux d'AdWords et de push Google shopping nettement plus réduits, de respectivement 15 % et 0 %. Cela témoigne de la faiblesse en référencement non naturel des titres qui ne font pas partie des principaux best-sellers.

c) L'autocomplétion

L'autocomplétion (ou complètement automatique) est une fonctionnalité introduite par Google en 2009 afin d'améliorer l'expérience utilisateur en suggérant les termes les plus fréquemment recherchés par les internautes à partir du ou des mot(s)-clé(s) demandés. En ce qui concerne les livres, cela induit en général l'ajout de mots-clés complémentaires comme le nom de l'auteur dans le cas d'une recherche par titre seul, ou le titre le plus populaire dans le cas d'une recherche par auteur.



D'après l'analyse de notre échantillon, les mots-clés les plus couramment suggérés sont « poche », « ebook », « gratuit », « résumé », « télécharger » et « pdf ».

A noter que Google, à la suite de pression des ayants-droit, a développé à l'automne 2011 un filtrage de l'autocomplétion en censurant les suggestions renvoyant vers des sites pirates, notamment de streaming et de torrent. Il conviendra donc d'observer l'évolution de l'application de la législation en ce domaine.

Annexes

Tableau n° 20 : Corpus n° 1 (tops *Livres-Hebdo* 2011)

Titre	Auteur	Editeur	Catégorie
La couleur des sentiments	Kathryn Stockett	J. Chambon	ROMAN
L'appel de l'ange	Guillaume Musso	XO	ROMAN
L'armée furieuse	Fred Vargas	V. Hamy	ROMAN
L'étrange voyage de monsieur Daldry	Marc Levy	R. Laffont	ROMAN
Rien ne s'oppose à la nuit	Delphine de Vigan	Lattès	ROMAN
Limonov	Emmanuel Carrère	P.O.L	ROMAN
L'art français de la guerre	Alexis Jenni	Gallimard	ROMAN
L'enfant allemand	Camilla Läckberg	Actes Sud	ROMAN
Le passager	Jean-Christophe Grangé	Albin Michel	ROMAN
1Q84	Haruki Murakami	Belfond	ROMAN
Twilight : Révélation	Stephenie Meyer	Hachette Jeunesse	JEUNESSE
Journal d'un vampire	Lisa Jane Smith	Hachette Jeunesse	JEUNESSE
Héros de l'Olympe	Rick Riordan	Albin Michel Jeunesse	JEUNESSE
Hunger games	Suzanne Collins	Pocket Jeunesse	JEUNESSE
Journal de Stefan	Lisa Jane Smith	Hachette Jeunesse	JEUNESSE
Percy Jackson	Rick Riordan	Albin Michel Jeunesse	JEUNESSE
Artemis Fowl	Eoin Colfer	Gallimard Jeunesse	JEUNESSE
Les chevaliers d'Emeraude	Anne Robillard	M. Lafon	JEUNESSE
Oksa Pollock	Anne Plichota	XO	JEUNESSE
Les héritiers d'Enkidiev	Anne Robillard	M. Lafon	JEUNESSE
Indignez-vous!	Stéphane Hessel	Indigène	ESSAI
On peut se dire au revoir plusieurs fois	David Servan-Schreiber	R. Laffont	ESSAI
Petit traité de vie intérieure	Frédéric Lenoir	Plon	ESSAI
Le chemin de l'espérance	Stéphane Hessel Edgar Morin	Fayard	ESSAI
Steve Jobs	Walter Isaacson	Lattès	ESSAI
De coeur inconnu	Charlotte Valendrey	Le Cherche Midi	ESSAI
Mémoires	Jacques Chirac	NIL	ESSAI
Engagez-vous!	Stéphane Hessel Gilles Vanderpooten	L'Aube	ESSAI
M. le Président	Franz-Olivier Giesbert	Flammarion	ESSAI
Sauvez votre argent!	Marc Fiorentino	R. Laffont	ESSAI
XIII Le jour du Mayflower	Sente Jigounov	Dargaud	BD
Thorgal Le bateau-sabre	Rosinski Sente	Le Lombard	BD
XIII mystery Colonel Amos	Alcante François Boucq	Dargaud	BD
Blake et Mortimer La porte d'Orphée	Van Hamme Aubin	Blake et Mortimer	BD
Les nombrils Un couple d'enfer	Delaf Maryse Dubuc	Dupuis	BD
Largo Winch Mer Noire	Van Hamme	Dupuis	BD
Boule et Bill A l'abordage!!	Laurent Verron	Dargaud	BD
Les Tuniques bleues Indien, mon frère	Cauvin Lambil	Dupuis	BD

Antarès Episode 4	Léo	Dargaud	BD
Julia & Roem	Enki Bilal	Casterman	BD
Nutella : le petit livre	Sandra Mahut	Marabout	PRATIQUE
Les recettes Dukan : mon régime en 350 recettes	Pierre Dukan	J'ai lu	PRATIQUE
La méthode Dukan illustrée	Pierre Dukan	Flammarion	PRATIQUE
Carambar : le petit livre	Garlone Bardel	Marabout	PRATIQUE
Les 100 aliments Dukan à volonté	Pierre Dukan	J'ai lu	PRATIQUE
Guide Rouge France 2011 + 2010		Michelin Cartes et Guides	PRATIQUE
Véritable petit beurre LU : les 30 recettes culte	Catherine Quévremeont	Marabout	PRATIQUE
Spéculoos Lotus	Brigitte Namour	Marabout	PRATIQUE
La pâtisserie Dukan	Pierre Dukan	J'ai lu	PRATIQUE
Crème de marrons C. Faugier : les 30 recettes cultes	Sandra Mahut Richard Boutin	Marabout	PRATIQUE

Tableau n° 21 : Corpus n° 2 (tops variés Datalib, Eyrolles.com, Didactibook.com et choix de titres *ad hoc*)

L'Empire Brisé	Mark Lawrence	Bragelonne	ROMAN
Laitier de nuit	Andrei Kourkov	Liana Levi	ROMAN
L'Invraisemblable histoire de Georges Pessant	Bertrand Leclair	Flammarion	ROMAN
L'Empreinte à Crusoé	Patrick Chamoiseau	Gallimard	ROMAN
Carénage	Sylvain Coher	Actes Sud	ROMAN
Promenade dans Berlin	Franz Hessel	L'Herne	ESSAI
Refuge sacré	Cathie Barreau	Publie.net	ESSAI
L'Inquiétude d'être au monde	Camille de Toledo	Verdier	ESSAI
Surfer la vie	Joël de Rosnay	Les Liens qui Libèrent	ESSAI
Pladoyer pour le bonheur	Mathieu Ricard	NIL	ESSAI
Le vinaigre	Nathalie Semenuik	Rustica	PRATIQUE
S'organiser tout simplement	Laurence Einfalt Stéphanie Bujon	Eyrolles	PRATIQUE
Petit Futé Barcelone 2012		Petit Futé	PRATIQUE
Comment ne plus se faire arnaquer par son banquier	Jean-François Pissard	Maxima	PRATIQUE
Les tarots clés en main	Lise Lebas	Fleur de Lys	PRATIQUE
Mademoiselle Princesse et Personne	Maryline Rich	Zabouille	JEUNESSE
Brousse Academy	Bernard Berger	La Brousse en Folie	JEUNESSE
Angélia : la Saga des Anges : Amour d'Enfer	J. A. Curtol	Les éditions Sharon Kena	JEUNESSE
La roue des vents	Vincent Joubert	Ankama Editions	JEUNESSE
Ava préfère les fantômes	Maïté Bernard	Syros	JEUNESSE
Portugal	Cyril Pedrosa	Dupuis	BD
La page blanche	Pénélope Bagieu	Delcourt	BD
Chroniques de Jérusalem	Guy Delisle	Delcourt	BD
Quai d'orsay	Blain Lanzac	Dargaud	BD
Un sac de billes Bailly	Bailly Kris	Futuropolis	BD

Tableau n° 22 : Top 50 des sites les mieux référencés sur Google (parmi les 10 premiers résultats, tous types de recherche confondus)

	Racine de l'URL	Occurrences*	%
1	livre.fnac.com	920	11,6%
2	amazon.fr	832	10,5%
3	decitre.fr	360	4,5%
4	fr.wikipedia.org	286	3,6%
5	ebooks-gratuit.com	113	1,4%
6	starzik.com	97	1,2%
7	lexpress.fr	88	1,1%
8	babelio.com	85	1,1%
9	evene.fr	84	1,1%
10	ebookbrowse.com	72	0,9%
11	itunes.apple.com	71	0,9%
12	fr-fr.facebook.com	67	0,8%
13	priceminister.com	62	0,8%
14	numilog.com	61	0,8%
15	t411.me	60	0,8%
16	didactibook.com	59	0,7%
17	librairie.immateriel.fr	58	0,7%
18	rueducommerce.fr	54	0,7%
19	filestube.com	51	0,6%
20	banque-pdf.fr	49	0,6%
21	booknode.com	48	0,6%
22	push Google	48	0,6%
23	nouvellebiblio.com	40	0,5%
24	fluket.com	38	0,5%
25	cultura.com	38	0,5%
26	bedetheque.com	36	0,5%
27	img.rechercher.me	34	0,4%
28	epagine.fr	34	0,4%
29	actuabd.com	34	0,4%
30	over-blog	34	0,4%
31	dargaud.com	29	0,4%
32	librairiedialogues.fr	29	0,4%
33	actualitte.com	28	0,4%
34	filesdrop.com	28	0,4%
35	thepiratebay.se	27	0,3%
36	telerama.fr	26	0,3%
37	lefigaro.fr	24	0,3%
38	viaouest.com	24	0,3%
39	franceculture.fr	23	0,3%
40	allocine.fr	22	0,3%
41	teamalexandriz.org	22	0,3%
42	telecharger-tout.com	21	0,3%
43	dupemonkey.com	20	0,3%

44	rechercher.me	20	0,3%
45	gallica.bnf.fr	20	0,3%
46	ebooks.telechargementgratuits.com	20	0,3%
47	bdgest.com	20	0,3%
48	planete-lolo.com	19	0,2%
49	123people.fr	19	0,2%
50	ciao.fr	18	0,2%

*Occurrences de liens en considérant les 10 premiers résultats de Google, sur un total de 7 913 liens différents.

Tableau n° 23 : Rangs moyens des 5 principaux sites référencés (tous types de recherche confondus ; cf. tableau n° 1, p.6)

Rang	Fnac	Amazon	Decitre	Wikipedia	Ebooks-gratuit
1	15 %	22,8 %	3,6 %	23,4 %	27,4 %
2	20,4 %	17,4 %	8,6 %	14,7 %	15 %
3	16,2 %	14,6 %	8 %	8,7 %	14,2 %
4	12,4 %	12,9 %	10,8 %	10,1 %	13,3 %
5	10 %	10,5 %	14,1 %	10,8 %	7,1 %
6	7,1 %	7,1 %	14,7 %	10,5 %	7,1 %
7	7,6 %	4,9 %	15 %	9,4 %	8 %
8	6 %	4,9 %	13,6 %	8,7 %	4,4 %
9	4,1 %	4,3 %	9,4 %	2,8 %	3,5 %
10	1,1 %	0,5 %	1,7 %	0,7 %	0 %
Total	100	100	100	100	100

Le tableau se lit ainsi : « 15% des liens Fnac.com issus d'une recherche sur les 10 premiers résultats de Google se positionnent au 1^{er} rang des résultats. »

www.lemotif.fr

En application de la réglementation en vigueur, il est interdit de reproduire cette étude intégralement ou partiellement, sur quelque support que ce soit, sans l'autorisation préalable écrite du MOTif ; cette étude ne peut faire l'objet d'aucune diffusion ou commerce sans l'autorisation préalable écrite du MOTif.