

EbookZ ?

Etude sur l'offre numérique illégale des livres français sur Internet en 2009

Mathias Daval | LE MOTif

Octobre 2009

Contacts

Mathias Daval
Edysseus Consulting
62 bd Diderot, 75012 Paris
www.edysseus.com
mathias.daval@edysseus.com

Cécile Moscovitz
LE MOTif
6 villa Marcel Lods
Passage de l'Atlas, 75019 Paris
www.lemotif.fr
cecile.moscovitz@lemotif.fr

Table des matières

Introduction.....	3
Champ de l'étude	5
L'offre numérique légale	6
Le piratage des livres aujourd'hui	7
Résumé des principales et récentes études.....	9
Quelques réflexions générales sur le piratage des livres	15
Panorama de la « scène <i>warez</i> »	19
Processus de piratage d'un livre papier	22
Méthodologie de l'étude.....	25
Les circuits de téléchargement illégal	25
Comment évaluer le nombre d'ouvrages disponibles illégalement et de téléchargements ?.....	30
Echantillonnage.....	36
Résultats.....	38
Analyse des fichiers	38
Analyse des ouvrages	42
Synthèse des résultats	52
Conclusion	53
Lexique	56
Annexes	57

Introduction

Avec plus de 76 000 titres produits en 2008 et près de 600 000 disponibles, l'édition française est extrêmement dynamique, même si son chiffre d'affaires a affiché un léger recul entre 2007 et 2008 (-2,2 % pour environ 2,83 Mds d'euros)¹.

La question du piratage s'inscrit dans la problématique de la protection du droit d'auteur (photocopillage, contrefaçon...), dans celle de la répartition des revenus entre les différents acteurs de la chaîne du livre, et plus largement dans l'avenir de l'édition à l'heure des bouleversements numériques.

Pour des raisons d'ordre culturel et technique, le piratage numérique est un enjeu difficile à cerner aujourd'hui par les éditeurs. Parfois attentistes, souvent vigilants, éditeurs et professionnels du livre partagent parfois un certain nombre de fantasmes qui ne reposent pas sur une observation concrète des échanges sur Internet.

Le problème posé par le téléchargement illégal est-il similaire à celui du photocopillage ? S'agit-il juste d'un changement d'échelle ? La démultiplication et l'immédiateté de la circulation des fichiers provoquent des réactions de peur à l'égard d'un marché parallèle : les pertes supposément infligées aux secteurs de la vidéo, de la musique et du logiciel doivent-elles faire craindre un sort similaire à l'industrie du livre ?

La présente étude est la première en France à analyser spécifiquement **l'offre illégale** des livres sur Internet, et à répondre notamment aux questions suivantes :

- Quelle est la nature de cette offre ?
- Quels sont les livres et les éditeurs les plus concernés ?
- Par quels circuits et comment sont piratés ces ouvrages ?
- De quelles quantités parle-t-on ?

Notre étude emploie le mot « piratage » dans le sens de : « copie et/ou diffusion de supports protégés par le droit d'auteur sans l'autorisation des ayants droit. » Le terme a été choisi pour des raisons pratiques, car il permet d'évoquer à la fois les internautes qui mettent à disposition les fichiers et ceux qui les téléchargent (que nous appellerons par la suite « téléchargeurs »). Il serait donc plus juste de parler de « mise à disposition et/ou de téléchargement illégal de livres au format numérique » que de piratage. Ce terme, dans son acception judiciairisée, évoque en effet davantage un phénomène de contrefaçon organisée dans un but lucratif, ce qui n'est pas le cas en matière de livres, comme en témoigne notre recherche pour les besoins de cette étude.

Celle-ci, comme nous l'avons dit, se concentre sur le premier des termes mentionnés ci-dessus : la mise à disposition. Car, en effet, avant de chercher à savoir ce qui est téléchargé, encore faut-il savoir ce qui est accessible. Or, ce type d'investigation n'a pas été mené de manière convaincante jusqu'à présent — nous revenons sur ce point p.10 à 15 dans la partie « Résumé des principales études ».

¹ Chiffre des ventes des éditeurs aux diffuseurs-distributeurs. Toutes exploitations confondues, ce chiffre s'élève à un peu plus de 5 Mds €. Source : SNE.

Après un rappel de quelques enjeux théoriques, économiques et pratiques du téléchargement illégal, nous nous pencherons sur ses circuits de diffusion ainsi que sur les processus de piratage numérique des livres. Nous continuerons avec le détail et l'analyse des résultats. Enfin, on se reportera aux annexes pour compléter les données chiffrées.

Champ de l'étude

Contenus

Le sujet de l'étude est le « piratage numérique du livre » et non le « piratage du livre numérique » : il n'est pas restreint aux seuls titres disponibles en offre numérique **légale**, mais bien à l'ensemble des textes piratés, dont l'essentiel provient de livres papier scannés et diffusés en ligne. Lorsque nous parlerons de « livres piratés », il s'agira donc de l'ensemble des caractéristiques de **l'offre numérique illégale**, qu'elle provienne du scannage de livres imprimés ou du *cracking* (« craquage ») de livres déjà existants au format numérique¹.

La notion de « livre numérique » ou **livrel** suppose une redéfinition juridique du mot « livre », jusqu'alors traditionnellement réservée aux imprimés non périodiques (« *Un livre est un ensemble imprimé, illustré ou non, publié sous un titre ayant pour objet la reproduction d'une œuvre de l'esprit d'un ou plusieurs auteurs en vue de l'enseignement, de la diffusion de la pensée et de la culture* », telle est la définition fiscale du livre), à l'exclusion donc de la presse².

Or l'usage courant du mot **ebook** en matière de piratage, que l'on retrouve aussi sous la forme « **ebookz** » ou « **bookwarez** », recouvre les livres mais aussi les périodiques. En pratique, ces derniers constituent même 70 à 80 % du nombre de fichiers disponibles en téléchargement illégal³. Nous les avons systématiquement exclus des fichiers analysés.

L'étude ne concerne que les **contenus payants et sous droits**, donc tombant sous le coup de la législation française sur le droit d'auteur. La mise à disposition gratuite de contenus numériques issus du domaine public, mais qui peuvent constituer une partie non négligeable du chiffre d'affaires de certains éditeurs, est un enjeu qui déborde le cadre de notre étude et qui concerne une quantité infime des échanges en matière de peer to peer*.

Enfin l'étude ne tiendra pas compte du piratage des traductions d'œuvres françaises. Elle ne concerne que les **livres publiés par des éditeurs français, en français** (traduites ou non).

Fichiers analysés

Nous avons défini des échantillons ou ensembles d'ouvrages disponibles au téléchargement illégal, divisé en 3 catégories : les **livrels**, les **bandes dessinées** (BD) et les **livres audio**, qui comportent respectivement environ 500, 300 et 100 titres. Ces ensembles ont été constitués par les résultats de recherches croisées sur différents circuits de diffusion et affinées en excluant les fichiers corrompus ou ne correspondant pas à des livres sous droit. (La méthode utilisée pour définir ces ensembles est détaillée p.38).

NB : Les termes de l'étude suivis d'un astérisque* sont définis dans le lexique p.58.

¹ La distinction entre ces deux origines fait l'objet d'une analyse spécifique (cf. p.40).

² Nous avons également exclu de l'étude les partitions musicales, qui font l'objet d'un piratage massif et spécifique.

³ D'après notre observation des rubriques « ebooks » des principaux *trackers** et sites de référence en peer to peer*. Cf. p.31-32.

L'offre numérique légale

La numérisation des œuvres est l'un des grands chantiers culturels sur Internet depuis une dizaine d'années. En témoigne la vitalité de projets comme Gutenberg Project aux États-Unis, Europeana en Europe et Gallica⁵ en France.

En parallèle de la diffusion de contenus gratuits et issus du domaine public, dont nous ne ferons pas cas dans notre étude, une offre payante a commencé à se développer à travers : a) les sites de certaines maisons d'édition (ex. : Eyrolles avec <http://izibook.eyrolles.com>), b) des agrégateurs indépendants (Immateriel, E-pagine) ou appartenant à des groupes : les deux principaux sont Numilog (racheté par Hachette en 2008) et Eden-Livres (lancé en septembre 2009 par un consortium d'éditeurs — Gallimard, Flammarion et La Martinière).

Le marché numérique existe depuis plusieurs années dans certains secteurs comme les éditions scientifiques, scolaires et pratiques. Aux États-Unis, le livrel connaît une forte croissance, avec près de 26 M \$ pour le 1^{er} trimestre 2009 contre 11,2 M \$ pour le 1^{er} trimestre 2008⁶, et des prévisions estimant à 3 % la part du numérique dans le marché du livre en 2009. Un éditeur américain comme Loose Id (www.loose-id.com), spécialisé dans les romans érotiques et sentimentaux à la manière des collections Harlequin, a réalisé un chiffre d'affaires d'1,3 M \$ en 2008 en ventes numériques.

En France, le marché de l'offre légale payante est encore extrêmement restreint. Selon le Syndicat National de l'Édition, le livre numérique représenterait en France de 30 à 40 millions €, soit un peu plus de 1 % du chiffre d'affaires de l'édition (2,83 Mds €), et l'essentiel des ventes étant celles de supports physiques de type CD/DVD⁷.

Au total, la marge de développement de l'offre numérique est énorme puisque, selon une étude de la Commission européenne⁸, près de 90 % des livres des bibliothèques de l'Union ne sont plus disponibles à la vente, qu'ils soient épuisés ou orphelins⁹. Même parmi les nouveautés, qui constituent l'essentiel des ventes en librairie, seule une partie est aujourd'hui achetable en version numérique. Encore embryonnaire dans la période de transition qui précède la numérisation de l'ensemble de ces fonds, la question du piratage pourrait donc devenir beaucoup plus essentielle dans les années à venir.

Principales données sur le marché du livre sur Internet en France (2009) :

Nombre d'internautes ¹⁰	34 M ^a
Nombre d'abonnements haut débit	17,7 M ^b
Ventes de livres en ligne (VPC) ¹¹	300 M € ^c
Ventes de livres numériques en ligne	30-40 M € ^c (dont 40 000 téléchargements sur Fnac.com ^d)
Ventes d'appareils nomades	45 M dont 2 M d'Iphones ^e
Ventes de liseuses	10 000 Sony Reader vendus par la Fnac en 2009 ^d

Sources : ^aForrester Research, ^bArcep, ^cGFK, ^dFnac, ^eGreenwich Consulting

⁵ <http://www.gutenberg.org>, <http://www.europeana.eu>, <http://gallica.bnf.fr>

⁶ Chiffre des ventes des éditeurs aux diffuseurs-distributeurs : www.idpf.org/doc_library/industrystats.htm

⁷ <http://www.sne.fr/pages/informations/livre-electronique-03-09.html>

⁸ <http://europa.eu/rapid/pressReleasesAction.do?reference=SPEECH/09/336&format=HTML&language=EN>

⁹ Un ouvrage est « orphelin » lorsqu'il n'a pas d'ayant-droit clairement reconnu, créant une situation dans laquelle il est délicat de négocier ses droits numériques.

¹⁰ 67 % des plus de 18 ans et 92 % des 12-17 ans ont un ordinateur chez eux (Source : Credoc, novembre 2008).

¹¹ A noter que, selon la Fevad, les ventes en ligne ont résisté à la crise en 2009 et ont même progressé, tous produits culturels confondus, de 25 % au 1^{er} trimestre par rapport à 2008.

Le piratage des livres aujourd'hui

« Sur la question des droits d'auteur et du piratage, les chiffres sont, par définition, très difficiles à évaluer. Il semblerait que le piratage de livres concerne un peu plus de 20 % des droits d'auteur servis chaque année – ce qui est considérable –, mais cette statistique doit être prise avec beaucoup de précautions. En outre, le piratage concerne vraisemblablement plus les sciences humaines que la littérature générale. »
(Hervé Gaymard, commission des affaires culturelles de l'Assemblée Nationale, séance du 25/03/2009)¹²

Ce chiffre de 20 %, selon le rapport Gaymard, est une évaluation sans référence précise qui concerne à la fois le piratage numérique et la contrefaçon en France. Il est à rapprocher des études annuelles de l'International Intellectual Property Alliance qui tente d'évaluer chaque année les pertes liées à la contrefaçon de livres¹³. En Chine et en Inde, on estime qu'un best-seller sur deux est contrefait en moins de 6 mois par des imprimeurs peu scrupuleux. Au Pérou, les bénéficiaires de l'industrie pirate dépassent ceux du marché légal. Là, pas de philosophie altruiste de partage de la culture, puisqu'il s'agit de véritables économies parallèles.

Mais si le piratage par réimpression illégale ou photocopie de livres papiers, en particulier dans les pays en développement, est un phénomène ancien dont on peut évaluer plus ou moins l'impact, le piratage numérique est une donnée nouvelle difficile à estimer. D'après sondage Ipsos relayé par l'AFP (18 septembre 2009), 15 % des Français avouent télécharger illégalement (44 % pour la moyenne mondiale). Beaucoup de généralités sont relayées par les médias mais aussi dans les rapports officiels ainsi que dans les discussions sur le téléchargement illégal de livres. Ainsi le rapport Gaymard¹⁴ souligne la part d'idées reçues en ce domaine (p.170) :

Marie-Pierre SANGOUARD

Je peux apporter un début de réponse. Tout d'abord, il ne faut pas se tromper : le téléchargement gratuit du livre sur internet existe déjà. Vous l'avez partout, vous avez des bandes dessinées intégralement téléchargeables. Pour ce qui est des sciences humaines, je ne vous en parle même pas.

Henri CAUSSE

Ça a commencé à se répandre, c'est vrai, mais ce sont le plus souvent de mauvais PDF ! C'est une opinion très largement répandue aujourd'hui, et qui est pourtant loin de la réalité (...). Une observation similaire peut être faite sur les catégories d'ebooks piratés, qui ne sont pas, contrairement aux idées reçues, réservées aux seuls ouvrages de STM et aux bandes dessinées, même si ceux-ci sont évidemment les plus répandus sur internet.*

Concernant certains genres comme les ouvrages dits de STM* (Sciences, technique et médecine) ou de bandes dessinées, l'offre illégale existe en effet depuis plus de dix ans sur Internet. Mais celle-ci ne saurait être limitée à ces secteurs.

¹² <http://www.assemblee-nationale.fr/13/cr-cafc/08-09/c0809047.asp>

¹³ <http://www.iipa.com/rbc/2009/2009SPEC301LOSSLEVEL.pdf>. Il est à noter que les estimations de l'IIPA sont plutôt basses, et que la profession y applique des coefficients multiplicateurs de l'ordre de 5 à 10 pour ses propres estimations.

¹⁴ http://www.culture.gouv.fr/culture/actualites/rapports/rapport_gaymard.pdf

Il convient également de souligner que chaque marché national (ou plutôt linguistique) est extrêmement spécifique. Ainsi le piratage numérique de livres est moindre dans l'Hexagone qu'en Italie ou qu'en Espagne. Certains best-sellers d'auteurs français sont même autant disponibles dans d'autres langues qu'en français¹⁵.

Enfin, juridiquement, la question du droit d'auteur est complexe, car la durée et les conditions de ces droits est variable selon les pays. Ainsi de nombreux textes français ont été diffusés sur des sites canadiens en vertu de la législation locale qui porte la protection à 50 ans après la mort de l'auteur et non à 70 ans. Ce qui permet à ces sites de diffuser par exemple en toute légalité, bien que leur téléchargement constitue un délit en France, des œuvres de Maurice Leblanc (mort en 1941). Notons que la contrefaçon couvre également les ouvrages issus du domaine public lorsqu'ils sont dotés de spécificités propres à une réédition moderne : préface, appareil critique, illustrations, etc. Le champ de notre étude n'inclut pas les ouvrages du domaine public, qu'ils soient ou non accompagnés d'éléments sous droits.

¹⁵ C'est le cas de *L'Élegance du hérisson* de Muriel Barbery, qui au moment de cette étude n'était disponible que dans une ou deux sources en français, contre plusieurs dizaines dans d'autres langues. Il n'est pas sûr, toutefois, que ces best-sellers français soient plus téléchargés que ceux des auteurs locaux.

Résumé des principales et récentes études évoquant le téléchargement illégal des livres

Il existe très peu d'études sur le piratage numérique des livres parce que, jusqu'à présent, il s'agit d'un marché parallèle encore très marginal comparé à celui de la musique, du film ou du jeu vidéo. Par ailleurs, comme toute évaluation d'une offre de nature souterraine et illégale, les méthodes d'investigation n'autorisent pas de conclusions définitives.

Les rapports du gouvernement (2008-2009)

Remis le 30 juin 2008 à Christine Albanel juste après la présentation du premier rapport sur le projet de loi Hadopi en conseil des Ministres, le **rapport Patino**¹⁶ ne fait que peu mention du téléchargement illégal. Il évoque juste le fait que les DRM* sont, lorsqu'ils gênent l'interopérabilité, contre-productifs dans la lutte contre le piratage. Il préconise plutôt le développement d'une offre légale attractive.

Mêmes conclusions de la part du **rapport Gaymard**, remis en mars 2009, qui comporte de courts paragraphes (p.106-108) sur la question du piratage, soulignant qu'il s'agit d'une « grande interrogation du secteur du livre à l'heure actuelle », particulièrement dans le secteur de la bande dessinée et dans une moindre mesure des sciences humaines. Le rapport rappelle, à propos de la loi Création et Liberté dite « Hadopi » :

« L'article 2 du projet de loi¹⁷, même s'il ne vise pas directement le secteur de l'édition et du livre, le concerne pourtant tout autant que les autres puisque, comme les ayants droit et syndicats professionnels des autres secteurs culturels, ceux du livre pourront saisir l'Hadopi en cas de constatation d'une infraction à la législation relative aux droits d'auteur. Pour ce faire, le secteur doit absolument s'entendre sur les moyens appropriés de constater ou faire constater ce piratage. À cet égard, une solution pourrait passer par un accord avec le CFC (Centre français d'exploitation du droit de copie), qui en tant que société de perception et de répartition des droits, dispose déjà d'agents assermentés. La filière devrait également, comme le secteur de la musique le fait déjà, travailler avec un prestataire technique commun qui détecterait les infractions. Par ailleurs, s'agissant des mesures techniques de protection, l'article 9 quater du projet de loi, issu d'un amendement sénatorial, prévoit la signature par les organisations du secteur musical d'un accord relatif à l'interopérabilité des fichiers musicaux et à la promotion d'une offre légale d'œuvres sans protection. Cet article légifère donc non pas sur un abandon global des mesures de protection mais sur la possibilité pour les internautes qui prennent la peine d'acheter légalement des contenus de pouvoir les lire sur tous les supports de leur choix. Cette préoccupation est au cœur du rapport sur le livre numérique de Bruno Patino. Un groupe de travail étudie actuellement, dans le cadre du Conseil du livre, les moyens de garantir l'interopérabilité des normes techniques et des matériels de lecture afin d'assurer aux utilisateurs la pérennité des livres numériques qu'ils acquerront. On ne peut que souhaiter que ce groupe de travail observe de près les décisions que la future loi amènera le secteur musical à prendre, pour s'en inspirer au besoin. En effet, faute d'une telle réflexion, et alors que les offres légales d'ouvrages en ligne se mettent

¹⁶ <http://www.culture.gouv.fr/culture/actualites/conferen/albanel/rapportpatino.pdf>

¹⁷ <http://www.culture.gouv.fr/culture/actualites/dossiers/internet-creation08/6%20-%20Projet%20de%20loi.pdf>

progressivement en place, comme dans le secteur de la musique, les internautes se tourneront vers l'offre piratée... »

Le rapport Equancy/Tera (novembre 2008)

Sous le titre : « Impact économique de la copie illégale des biens numérisés en France », l'étude réalisée par les cabinets de conseil Tera Consultants et Equancy & Co¹⁸ en novembre 2008 a été l'une des premières à chiffrer le préjudice du piratage sur le secteur de l'édition. Celui-ci est estimé à 147 millions d'euros (soit environ 5 % du chiffre d'affaires du secteur), essentiellement concentré sur la littérature scientifique et scolaire, et équivalant selon l'étude à 733 emplois directs à temps plein perdus.

La méthode prospective pour parvenir à ces conclusions est la suivante :

- 1) Elle part du nombre d'internautes en France (31,57 M),
- 2) Elle évalue ensuite combien de ces internautes utilisent le peer to peer* (9,5 M),
- 3) Elle adopte alors une première hypothèse : « *On fait donc l'hypothèse que 50% des usagers P2P [peer to peer*] lisent des formats numériques* » (4,75 M),
- 4) Elle multiplie le nombre d'internautes lisant illégalement au format numérique par le nombre de livres moyens achetés par les Français (7,29),
- 5) Elle applique au résultat trouvé (34,5 M) un taux de substitution des achats par le téléchargement illégal de 50 %. Ce qui donne donc les 17,25 M de titres piratés par an, qu'il suffit de multiplier par une estimation du prix du livre numérique (8,5 €) pour obtenir les 147 M € de perte pour le secteur du livre.

Une telle méthode pose problème car elle repose sur un taux de substitution purement arbitraire. Aucune étude ne vient étayer ce choix, pourtant essentiel car c'est lui qui donne la mesure des conséquences du piratage. Par ailleurs, la catégorie ciblée est uniquement celle des téléchargeurs « adeptes de lecture numérique »¹⁹, et l'étude considère que cette substitution s'opère d'un livre numérique payant à un livre numérique pirate, et non pas depuis un livre imprimé papier. Enfin, elle présuppose qu'un livre téléchargé correspond à une vente perdue, une affirmation que plusieurs études dans le monde de la musique ont déjà remise en cause²⁰.

On peut, à juste titre, considérer que l'utilisation de tels postulats est nécessaire pour parvenir à des évaluations prospectives. Ce qui est regrettable, en revanche, est que les résultats chiffrés, qui n'étaient que de simples hypothèses, ont été utilisés de façon univoque lors de la promotion de la loi Hadopi au début de l'année 2009. Rien ne prouve que les pertes de chiffre d'affaire comme d'emplois soient réelles. La corrélation entre le piratage et l'évolution économique du secteur du livre reste à démontrer.

¹⁸ <http://www.teraconsultants.fr>, <http://www.equancy.com>

¹⁹ « *Considérant que ces ventes perdues impactent uniquement la distribution numérique du livre (la catégorie ciblée est celle des téléchargeurs adeptes de lecture numérique, ils substituent donc le téléchargement à l'achat d'un livre numérique), quelle est la valeur d'un livre numérique ?* » (p.61)

²⁰ Cf. études détaillées ci-après.

Figure 56 : Estimation du nombre de ventes de livres perdues à cause de la copie illégale

# livres vendus en 2006 (Munités)	469,70
population française (M habitants)	64,47
# livres achetés par personne par an	7,29
# internautes (M internautes)	31,57
% usagers P2P chez les internautes	30%
pénétration de l'e-book chez les usagers P2P	50%
# livres achetés par usagers P2P détenteur d'un e-book (M unités)	34,50
% de téléchargement gratuit remplaçant l'achat	50%
# de téléchargement gratuits remplaçant l'achat (M unités)	17,25

Source : Analyse TERA Consultants

Les études de TNO (Pays-Bas, février 2009) et Oberholzer-Gee & Strumpf (États-Unis, mai 2009)

Une étude menée par le cabinet TNO²¹ aux Pays-Bas pour le compte de plusieurs ministères et publiée en février 2009, remet précisément en cause la corrélation entre piratage et baisse du chiffre d'affaires des industries culturelles. Intitulée « Economic and cultural effects of file sharing on music, film and games », l'étude va même jusqu'à souligner les conséquences **positives du partage de fichiers sur le marché néerlandais à court et à long terme**. Elle montre en effet que ceux qui téléchargent le plus sont aussi ceux qui achètent le plus de produits culturels. L'étude ne traite pas le marché des livres mais pose la question, au regard des conclusions trouvées pour tous les autres biens culturels numériques : « *Comment les éditeurs se débrouilleront à l'avenir avec l'apparition des ebooks ?* » (p.123 ; nous traduisons).

Les remarques sont similaires concernant la nouvelle étude²² « File Sharing and Copyright » menée par les deux chercheurs américains Oberholzer-Gee et Strumpf (Harvard Business School), déjà auteurs de travaux de recherche sur l'impact du peer to peer* en 2005 et 2007. Ils affirment que le piratage n'a pas fait baisser la demande de produits culturels, et qu'il faut briser le mythe : un fichier téléchargé = une vente perdue. « *Il n'y a aucune preuve que le partage de fichiers ait réduit la production d'œuvres artistiques* » (nous traduisons), affirment-ils.

Là encore, mais dans un sens opposé à l'étude Equancy/Tera, il convient de relativiser ces conclusions. Si rien ne prouve l'impact négatif du téléchargement illégal de contenus culturels, rien ne prouve non plus son impact positif. A l'aube d'un nouveau terrain de piratage numérique, celui des livres, une prudence élémentaire s'impose.

²¹ http://tno.nl/content.cfm?context=markten&content=publicatie&laag1=182&laag2=1&item_id=473

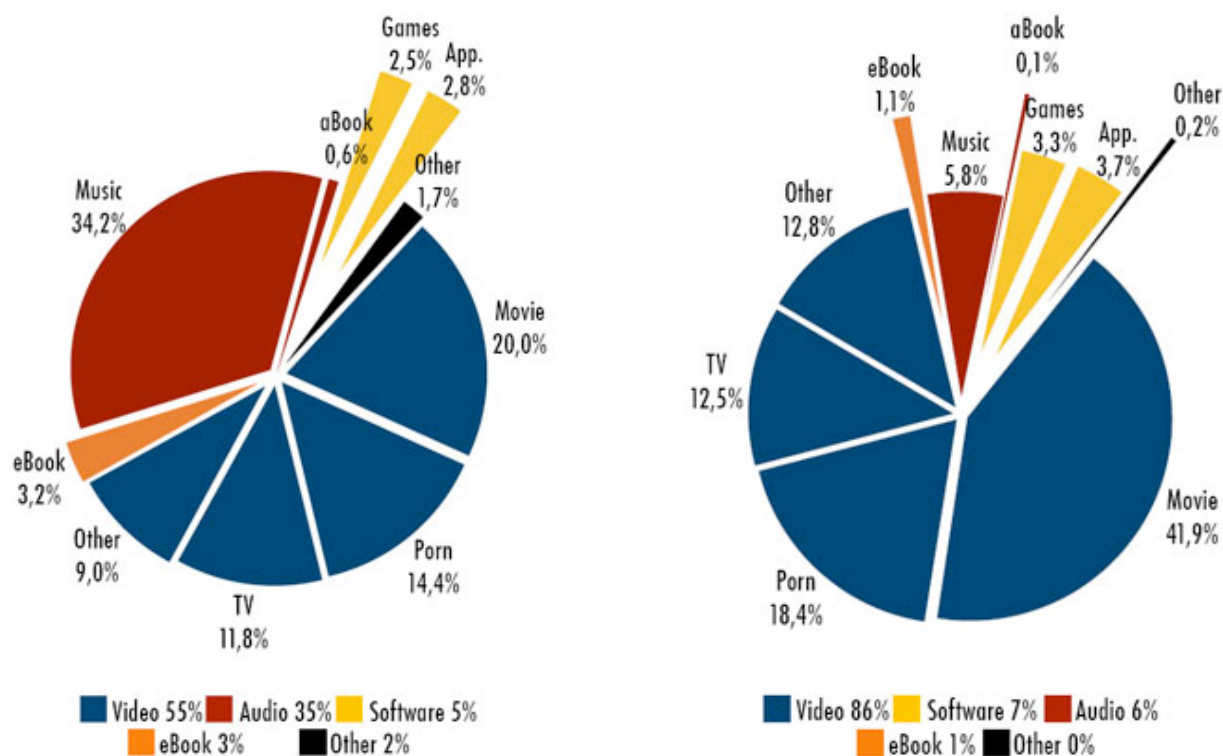
²² <http://www.hbs.edu/research/pdf/09-132.pdf>

Le cabinet de conseil allemand Ipoque a réalisé sa troisième étude annuelle²³ consacrée aux échanges sur Internet, qui analyse les usages du P2P* dans différents pays à travers le monde. C'est l'une des rares études où les livrels (*eBooks*) et les livres audio (*aBooks*) figurent dans les catégories de fichiers téléchargés.

Si les livres numériques sont encore absents des réseaux BitTorrent* (ou du moins échangés en quantités trop faibles pour figurer dans les résultats de l'étude), en revanche ils apparaissent dans les échanges par eDonkey*, comme le montre les diagrammes ci-dessous consacrés à l'Europe du Sud : 1,1 % du trafic (mesuré en mégaoctets) pour les livrels et 0,1 % pour les livres audio ; soit en nombre de fichiers uniques 3,2 % pour les livrels et 0,6 % pour les livres audio.

Les différences entre trafic et nombre de fichiers mises en évidence par les deux diagrammes sont expliquées par la taille moyenne des fichiers : on observe un ratio de 1:3 pour le livre²⁴, pour lesquels il y a une multitude de sources de petite taille (en général inférieur à 20 Mo), en revanche un ratio de 2:1 pour le film où les sources sont plus restreintes et généralement de taille supérieure à 600 Mo.

Dans son étude 2007, Ipoque ajoute en annexe les classements des livrels les plus téléchargés selon le mode d'échanges peer to peer* et la zone géographique (cf. annexe 1).



Répartition des contenus sur eDonkey (Europe du Sud), en nombre de fichiers uniques. Source : Ipoque.
 Répartition des contenus sur eDonkey (Europe du Sud), en trafic de téléchargement. Source : Ipoque.

²³ http://www.ipoque.com/resources/internet-studies/internet-study-2008_2009

²⁴ C'est-à-dire que les livrels disponibles illégalement représentent, en pourcentage de l'ensemble des contenus piratés en P2P*, trois fois plus en nombre de titres qu'en mégaoctets.

Etude du M@rsouin (2008)

Le Môle Armoricaïn de Recherche sur la Société de l'Information et les Usages d'INternet a publié l'année dernière une étude sur les pratiques de consommation des internautes en contenus illégaux²⁵. L'étude apporte une donnée concernant les livres, dans un tableau sur la répartition des types de biens culturels téléchargés, concernant à la fois téléchargements légaux et illégaux :

De la musique	77 %
Des logiciels	49 %
Des films	41 %
Des séries TV	23 %
Des documentaires	15 %
Des événements sportifs	9 %
Des livres	6 %

Répartition des téléchargeurs. Source : M@rsouin.

On y constate que seulement 6 % des téléchargeurs s'intéressent aux livres, soit plus de 10 fois moins que la musique et 7 fois moins que les films. Elle montre donc combien le piratage des livres est un phénomène encore marginal comparé aux autres formes d'échanges en peer to peer*.

Etude du SNE/GFK : « Tendances du marché du livre 2008-2009 » (mars 2009)

L'étude présentée au Salon du Livre par le Syndicat National de l'Édition indique que le livre numérique est « *un paradoxe commercial* » : « *Alors que certains éditeurs du secteur de l'information scientifique, technique et juridique réalisent d'ores et déjà 50 % de leur chiffre d'affaires mondial en ventes de contenus numériques et que le recul des ventes d'encyclopédies, dictionnaires ou cartes de géographie est en partie dû à une substitution vers le contenu numérique, les ventes de livres numériques réalisées en France par des éditeurs de livres comptabilisées dans l'enquête annuelle de branche, restent confidentielles depuis 3 ans, comprises entre 30 et 40 millions d'euros (soit 1 % du CA éditorial), essentiellement sur support tangible (CD/DVD).* »²⁶

Cette constatation présente la réalité du marché numérique aujourd'hui, à mettre en regard des enjeux du piratage, dont le rapport ne parle pas directement.

L'expérience menée par Magellan Media (2009)

Présentées en février 2009 à l'occasion du Tool of Change for Publishing Conference organisée par l'éditeur américain O'Reilly, deux expériences menées en 2008 par le cabinet de conseil Magellan Media ont étudié l'impact direct de la diffusion de livres sur le web²⁷ :

- D'une part, l'éditeur Random House a mis à disposition quelques titres gratuitement sur son site web : les ventes papier de ces titres ont augmenté en moyenne de 19,1 % pendant la période promotionnelle (par rapport aux ventes pendant les 4 semaines qui la précèdent), puis de 6,5 % pendant la période suivante. Notre étude montre que, si l'on ne peut en tirer de

²⁵ http://www.marsouin.org/article.php3?id_article=264

²⁶ http://www.sne.fr/pdf/Tendances_activite_editoriale2008-2009.pdf

²⁷ <http://oreilly.com/catalog/9780596157876/> et <http://www.magellanmediapartners.com>

causalité stricte quant à l'impact d'une diffusion gratuite sur les ventes (puisqu'on ne spéculer sur le montant des ventes si cette opération promotionnelle n'avait pas eu lieu), en revanche elle peut affirmer que cet impact n'est en aucune façon négatif.

- D'autre part O'Reilly a suivi certains de ses best-sellers, vendus sans DRM* sur son site. Il a observé la courbe de diffusion illégale de 8 titres piratés, et a constaté que leurs ventes numériques ont augmenté globalement de 6,5 % dans les 4 semaines qui ont suivi la première apparition de *seeds** (sources de fichiers illégaux) sur Internet. Par ailleurs, l'éditeur a été surpris de constater le temps de latence existant entre la publication officielle et la diffusion pirate : 20 semaines en moyenne (avec un écart allant de 4 à 43 semaines selon les titres).

L'étude reste toutefois extrêmement prudente et ne tire pas de conclusions quant à une corrélation directe entre ces différents facteurs. Il observe que la grande majorité des internautes qui téléchargent les livres illégalement sont ceux-là mêmes qui en parlent dans toutes les communautés en ligne (Facebook, Twitter, blogs, etc.) et qui en font en quelque sorte de la publicité gratuite. Cela va dans le sens de l'étude réalisée par deux chercheurs de l'université de Yale en 2005, qui montre l'impact (positif comme négatif) sur les ventes des livres en ligne des critiques déposées par les visiteurs d'Amazon.com²⁸. Les gros éditeurs américains ont depuis deux ans compris cet enjeu et développé des sites communautaires d'e-lecteurs, à l'image de Pulse It de Simon & Shuster ou de Spine Breakers de Penguin²⁹.

Une expérience similaire a été menée par un autre éditeur américain d'ouvrages informatiques, No Starch, en mars 2008³⁰. Un an plus tard, il n'est parvenu à aucune conclusion précise quant au lien entre la mise à disposition gratuite de leurs livres sur le web et l'évolution de leurs ventes. Ils sont en tout cas incapables d'y trouver un lien de causalité négatif, et ils demeurent partisans de cette méthode qui, selon eux, est efficace pour contrer le « bruit » généré par les milliers d'ouvrages qui sortent chaque année en librairie³¹. L'idée est qu'un éditeur a tout à gagner à se faire connaître ainsi par le bouche à oreille sur Internet.

Brian O'Leary, consultant de Magellan Media, est aujourd'hui en discussion avec d'autres éditeurs américains, dont Thomas Nelson, pour mener sur leurs catalogues des expériences similaires à celle d'O'Reilly. Il affirme : « *L'impact du piratage commence à devenir un problème pour les éditeurs, les auteurs et leurs agents. C'est maintenant qu'il faut réaliser des études afin d'obtenir des éléments de comparaison pertinents pour les années à venir, lorsque le marché numérique explosera.* »³²

Cet avis est partagé par Jack Shafer du magazine *Slate* dans son édition du 15/07/2009³³ : « *Aujourd'hui, le marché du livre électronique se trouve plus ou moins dans la même situation que le marché du mp3 en 1999, un an après la sortie du premier baladeur mp3. (...) Evidemment, il est facile de forcer le parallèle entre l'industrie du livre et de la musique, mais ce n'est qu'une question d'échelle. Chaque lecteur d'ebook qui décide de court-circuiter le système en téléchargeant des fichiers illégaux est un promoteur en puissance du marché noir du livre.* »

²⁸ <http://www.som.yale.edu/faculty/dm324/EffectWOMSalesdraftSep26.pdf>

²⁹ <http://pulseit.simonandschuster.com> et <http://www.spinebreakers.co.uk>

³⁰ <http://nostarch.com/blog/?p=127>

³¹ <http://nostarch.com/blog/2009/05/how-do-free-online-books-affect-sales/>

³² Entretien avec Brian O'Leary du 29/06/2009 (notre traduction).

³³ Traduit de : *Does the Book Industry Want to Get Napstereo?*
(<http://www.slate.com/id/2222941/pagenum/all/#p2>)

Quelques réflexions générales sur le piratage des livres

Le piratage des livres n'est pas une affaire nouvelle, mais l'enjeu était jusqu'alors réduit à la contrefaçon d'exemplaires papier et au « photocopillage » (lequel est l'objet du CFC, ou Centre français d'exploitation du droit de copie, c'est-à-dire la société de gestion collective pour la reproduction par reprographie de la presse et du livre), notamment dans les milieux scolaire et universitaire. Dans certains cas particuliers célèbres, ce sont même les éditions pirates qui ont contribué à faire connaître ou à diffuser les œuvres : ainsi l'édition non autorisée du *Seigneur des Anneaux* de J.R.R. Tolkien, publiée en 1965 par Ace Books aux États-Unis et répandue dans les campus universitaires, aidera l'édition légale chez Ballantine Books à devenir un best-seller dès l'année suivante.

Les différents points de vue des acteurs de l'édition

Mais l'entrée dans l'ère numérique pose le problème de façon beaucoup plus cruciale. Schématiquement, face au piratage des œuvres numériques, deux clans semblent s'affronter : d'un côté, les défenseurs de l'industrie culturelle et les défenseurs des droits d'auteur. De l'autre, les partisans (groupe pour le moins hétérogène) de la gratuité et du libre partage des œuvres. En réalité, et en dépit du clivage apparent dans les débats autour de la « loi Hadopi », le spectre est beaucoup plus large. Les médias ont eu tendance à réduire les enjeux, par « technomorphisme », à une simple opposition binaire inspirée des 1 et des 0 du numérique. C'est, à un extrême, Denis Olivennes³⁴ qui conteste le mythe de la gratuité et oppose au peer to peer* la survie de tout un pan de l'économie. Corollairement, pour Olivier Bomsel³⁵, le problème ne vient pas tant du piratage que des fournisseurs d'accès, ces intermédiaires qui se sont rendus incontournables et sont les véritables bénéficiaires de l'illusion du « tout gratuit ». A l'opposé, Hakim Bey³⁶ est porte-parole d'un mouvement libertaire dont la dimension numérique prône rien de moins qu'un nouveau modèle de société.

A quelques exceptions près qui ont monopolisé le débat médiatique dans un camp comme dans l'autre, on n'a que peu entendu les auteurs s'exprimer sur le piratage, souvent par méconnaissance des enjeux et de la culture numériques. Parmi ceux qui ont occupé le devant de la scène, Paulo Coelho a remarqué que dans les pays dans lesquels circulaient des versions pirates traduites par ses lecteurs, les ventes n'étaient pas affectées (et même avaient tendance à augmenter). Devant ce constat, cet auteur de best-sellers a franchi le pas en 2008 en diffusant lui-même ses ouvrages sur les réseaux peer to peer*³⁷.

Du côté des éditeurs, si beaucoup s'insurgent contre les dérives numériques³⁸, d'autres soulignent que les études sur l'impact du piratage restent relativement prudentes. Certains observateurs du secteur

³⁴ Ancien directeur de la Fnac, Denis Olivennes est le coordinateur éponyme du rapport remis à Christine Albanel en novembre 2007, à l'origine du projet de loi Hadopi. Il est l'auteur de *La Gratuité, c'est le vol* (Grasset, 2007).

³⁵ « *Les industriels n'hésitent pas à piller une industrie voisine pour relever l'utilité de leurs services* » (Neteco.com, 28/02/07). Olivier Bomsel est économiste et fut membre de la commission Olivennes. Il est l'auteur de *Gratuit !* (Folio, 2007).

³⁶ Ecrivain et théoricien politique américain, connu pour ses théories sur les TAZ ou « Temporary Autonomous Zones », dont la version numérique est la forme moderne des utopies pirates.

³⁷ « *Le piratage aide les œuvres à se vendre. Internet, c'est la démocratisation de l'information, et contrairement à ce que l'on pourrait penser, c'est le retour à l'écriture* » (<http://www.lewebparis.com/speakers.html>)

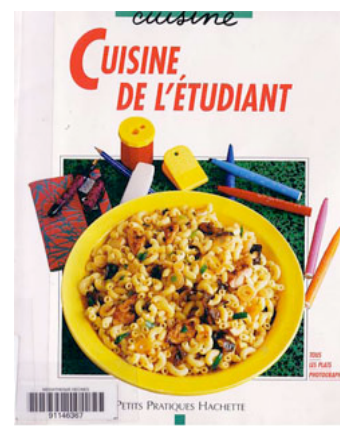
³⁸ John Wiley & Sons, un éditeur américain d'ouvrages techniques et universitaires (et de la série *Pour les Nuls*), emploie trois salariés à temps plein pour surveiller le piratage de ses livres. Source : <http://www.nytimes.com/2009/05/12/technology/internet/12digital.html>

du livre vont jusqu'à affirmer, non sans provocation, que c'est le piratage qui fera décoller les ventes des liseuses (*readers**), puisqu'il faut bien du contenu pour les alimenter et que le contenu légal est encore trop cher et réduit³⁹. C'est ce qui s'est passé dans le marché de la musique, où les premiers baladeurs mp3 étaient disponibles à la vente alors que l'offre légale n'existait pas encore. Il est certain que l'augmentation du piratage de livres traduit l'augmentation du besoin en offre numérique. Comme le rappelle le compte rendu des derniers États Généraux de la BD (juin 2009), les éditeurs ont l'impression d'être pris en étau : « *Partir bille en tête et risquer d'essayer les plâtres ou prendre le temps et se retrouver à la traîne ? Le choix n'est pas des plus simples*⁴⁰...»

C'est aussi ce que souligne l'éditeur américain Tim O'Reilly en considérant que, du point de vue des éditeurs comme des auteurs, le véritable problème n'est pas le piratage, mais l'obscurité⁴¹ : comment faire sortir un ouvrage de la masse sans cesse grandissante des productions éditoriales ? Cory Doctorow, essayiste, auteur de science-fiction et co-créateur du blog Boing Boing, apporte un argument supplémentaire : « *J'ai commencé à distribuer gratuitement mes livres lorsque j'ai été témoin des débuts de la scène « bookwarez ». (...) [Les] fans passaient facilement 80 heures pour pirater leurs livres favoris (...) J'ai considéré que c'était plutôt une bonne chose d'avoir à ma disposition 80 heures d'effort promotionnel gratuit*⁴². »

Le « trou analogique » : les problématiques technologiques de lutte contre le piratage

Un autre point important à souligner est le concept de « trou analogique » (« *analog hole* » en anglais)⁴³. Il indique le fait que toute œuvre, à partir du moment où on y a accès de façon analogique, peut être reproduite quelle que soient les mesures de protection sur sa version numérique. Ainsi un disque pourra toujours être reproduit directement depuis un poste de radio ou une version diffusée depuis une chaîne hi-fi. La particularité du livre par rapport à la musique ou la vidéo est le fait que le piratage passe aujourd'hui essentiellement par ce « trou analogique ». C'est-à-dire que les fichiers sources sont presque tous des livres papier, scannés et convertis dans le format adéquat⁴⁴. Le piratage de la musique, en revanche, a été dès l'origine immédiatement numérique puisqu'il s'agissait seulement de « ripper » la musique (convertir les informations contenues sur un disque ou un DVD sur son ordinateur), et que la protection des CD a rendu cette pratique un peu plus complexe. Ce qui veut dire que, en dépit des DRM* et des protections que pourront contenir les futurs livres numériques de l'offre légale, rien n'empêchera les pirates de continuer, comme traditionnellement, d'utiliser directement le livre papier. Et ce n'est pas l'accès à ces ouvrages qui pose problème, car une quantité non négligeable de livres piratés proviennent de bibliothèques (certains sont même scannés avec le code-barres, cf. visuel ci-contre).



³⁹ « Rampant Piracy Will Be The Kindle DX's Savior » (<http://www.techcrunch.com/2009/05/09/rampant-piracy-will-be-the-kindle-dxs-savior>)

⁴⁰ <http://syndicatbd.org/pdf/rapportbd.pdf>

⁴¹ « *Obscurity is a far greater threat to authors and creative artists than piracy.* » (<http://tim.oreilly.com/pub/a/p2p/2002/12/11/piracy.html>)

⁴² Traduit de « Free(konomic) Ebooks », in *Locus*, septembre 2007.

⁴³ <http://toc.oreilly.com/2008/10/the-analog-hole-in-digital-boo.html>

⁴⁴ Cf. p. 23.

Vers une redéfinition de l'économie de la diffusion du livre

Mais le centre du débat sur le livre numérique reste celui du prix. A l'offre payante semble s'opposer la gratuité du téléchargement illégal. Pourtant, pour Kevin Kelly, consultant spécialisé en nouvelles technologies et ancien rédacteur en chef de *Wired Magazine*, la gratuité des contenus culturels numériques ne fait pas tout⁴⁵. Il y a selon lui 8 autres facteurs dont les internautes tiennent compte et qui peuvent jouer en la faveur des offres légales : l'immédiateté, la personnalisation, l'interprétation, l'authenticité, l'accessibilité, l'incarnation, le mécénat, la trouvabilité. Ces 8 caractéristiques, contrairement à un simple fichier, ne peuvent pas être copiées. Comment ces caractéristiques s'appliquent-elles au téléchargement de livres ?

- **L'immédiateté** (*immediacy*) : télécharger un ouvrage illégal est encore une pratique réservée à des internautes avertis. Et, même si la maturité des usages va atténuer cette difficulté, il n'en restera pas moins que l'internaute est prêt à payer pour la facilité et la rapidité d'utilisation des sites légaux : le temps de recherche d'un fichier pirate est aussi un coût non négligeable.
- **La personnalisation** (*personalization*) : l'exemple de la musique montre qu'un consommateur télécharge peut-être gratuitement un morceau de musique, mais sera prêt à payer cher pour obtenir ce même morceau s'il est de très bonne qualité. C'est aujourd'hui déjà le cas pour certains types de musique comme la musique classique, pour laquelle la différence de qualité liée à la compression des mp3 est évidente : le véritable amateur ne peut se contenter d'un fichier mal compressé et se tournera plus facilement vers l'offre légale. Dans le cas du livre, il peut s'agir de proposer une offre personnalisée en fonction des goûts du lecteur (par exemple dans le cadre de réseaux sociaux sur le web).
- **L'interprétation** (*interpretation*) : c'est une logique qui provient du monde informatique où l'on dit souvent en plaisantant que le logiciel est gratuit mais que son mode d'emploi coûte 10 000 dollars. Il est possible d'imaginer une logique similaire dans le monde du livre technique ou pratique par exemple, où ce qui sera valorisé ne sera pas le livre lui-même mais ses applications : par exemple des cours de jardinage en ligne donnés par l'auteur d'un livre sur la botanique, etc.
- **L'authenticité** (*authenticity*) : le lecteur a besoin d'être sûr que le livre qu'il est en train de consulter est bien l'ouvrage authentique tel qu'il a été créé par son auteur. Qu'il n'y a pas d'erreurs typographiques ou orthographiques liées à une mauvaise saisie lors du scannage du livre, ou que des morceaux entiers n'ont pas été supprimés lors de la numérisation, ce qui arrive parfois. Et le risque de déformation ou de manipulation est véritable : autant le risque était compliqué et coûteux en vidéo, presque inutile en musique, autant pour le livre c'est extrêmement facile. Imaginons une secte modifier un texte religieux ou un même un roman pour y glisser ses messages. Idem pour les traductions non autorisées : on pense ici à celle d'*Harry Potter*, truffée d'erreurs et d'approximations.
- **L'accessibilité** (*accessibility*) : avec la multitude des fichiers et des terminaux à notre disposition, leur accessibilité devient de plus en plus complexe. Il est indispensable qu'il y ait de nouveaux services et outils. Ce qui va être valorisé finalement n'est pas tant la gratuité

⁴⁵ http://www.kk.org/thetechnium/archives/2008/01/better_than_fre.php

des fichiers que la possibilité de les stocker, de les sauvegarder, de les organiser de façon optimale, et tout cela a une valeur marchande.

- **L'incarnation** (*embodiment*) : avoir un produit virtuel gratuit, c'est bien, mais sa version « réelle » sera toujours mieux valorisée par le consommateur. Kevin Kelly donne l'exemple du fan qui va à un concert des Rolling Stones plutôt que de télécharger un de leurs mp3. Dans le secteur du livre, cela peut passer par une multitude d'actions qui ne peuvent avoir lieu que sur le territoire local, et qui sont extrêmement valorisées par le lecteur : « the book is free, the bodily talk is expensive » (lire un livre est gratuit mais écouter quelqu'un parler vaut cher).
- **Le mécénat** (*patronage*) : d'après Kevin Kelly, les lecteurs ont souvent envie de soutenir les créateurs. On peut imaginer mettre à disposition gratuitement un livre numérique avec la possibilité de donner en ligne la somme que l'on veut. Bien sûr le lecteur veut être certain que l'argent ira bien à l'auteur. Ce n'est pas un modèle économique suffisant en tant que tel, mais une alternative ou un complément⁴⁶.
- **La trouvabilité** (*findability*) : dans un environnement où des milliers (et bientôt des millions) de livres sont disponibles gratuitement, le lecteur veut qu'on lui propose des choix, que l'on opère des sélections pour lui. Le rôle des prescripteurs reste essentiel, qu'on les appelle des éditeurs, des agrégateurs, des diffuseurs ou des distributeurs.

L'ensemble de ces caractéristiques tend à faire valoir le fait que l'on passe d'une industrie de produits culturels (les livres) à une industrie de services culturels (les services liés aux livres).

Cette évolution ne modifie pas nécessairement la chaîne de valeur du livre et le rôle de chaque intervenant (auteur, éditeur, distributeur/diffuseur, libraire). Mais elle suppose une adaptation de chacun de ses acteurs à l'environnement numérique. Elle suppose également, ce que ne précise pas Kevin Kelly bien que ce soit toute la difficulté aujourd'hui, de définir un ou plusieurs modèles économiques concrets. Le secteur du livre n'est pas sorti, à cet égard, de sa phase de défrichage.

⁴⁶ A cet égard, l'exemple le plus marquant est celui du groupe Radiohead, qui a diffusé en 2007 son album *In Rainbows* entièrement en téléchargement, sans aucune autre contrepartie financière que le don.

Panorama de la « scène warez »

La création et la mise à disposition de fichiers numériques piratés est produite, en partie, par une nébuleuse d'intervenants généralement appelée « la Scène ». Rappelons que le mot *warez** signifie tout contenu protégé et diffusé illégalement sur les réseaux. La Scène livrels⁴⁷ est nettement moins développée que ses alter ego en musique ou en vidéo. Elle est néanmoins présente depuis la fin des années 1990, y compris dans le monde francophone⁴⁸. Elle regroupe des individus et des organisations extrêmement différents les uns des autres, qui vont de l'internaute moyen uploadant lui-même quelques fichiers de temps en temps sur un forum spécialisé, aux équipes (« *teams** ») les plus secrètes et organisées.

Dans l'univers du *warez**, le mot « ebook » ou « ebookz » a un sens un peu différent de l'acception commune, car, comme nous l'avons déjà évoqué, il inclut également la presse. Journaux et magazines constituent même la grande majorité des fichiers diffusés sur les différents réseaux⁴⁹. Il est par ailleurs difficile de suivre l'actualité du piratage des livres car ces derniers ne sont pas encore intégrés dans la plupart des sites d'informations sur les *releases** (aussi appelés « *release logs* » ou « *dupecheck* »), qui listent chaque jour les dernières diffusions⁵⁰.

Depuis quelques années, une demi-douzaine d'équipes ou « *release groups* » francophones se partagent les *releases** de qualité. Mais la Scène est par définition très volatile et certaines équipes disparaissent du jour au lendemain. Plus étonnante, et méconnue du grand public, la concurrence entre les équipes pour sortir les *releases** le plus rapidement possible et de la meilleure qualité qu'il soit, à l'image de cette annonce postée sur un fichier d'information .nfo⁵¹ :

« Notes a notre nouveau concurrent : On tiens a dire qu'on est un peu rageux de voir encore une nouvelle team FR sortir des rlz de magazines en soit disant "e-book" alors que comme les autres ce ne sont que des scans JPG. Les scans doivent etre retraites, recadres, retouches, mis en pdf et rendu interactif par l'ajout de liens. Ah oui forcement ca prend plus de temps. Mais il faudrait voir a rehausser un peu le niveau de vos rlz et aux site op qui acceptent ca d'etre un peu plus rigoureux. Pour info les releases e-book sont en pdf avec des liens (et meme en OCR sur la scene US). Rien avoir avec de vulgaires scan en JPG qui ne beneficent d'aucun retraitement de votre part. Allez un petit effort... » [sic]

La qualité des fichiers est donc vue comme un principe essentiel, même s'il n'est pas toujours respecté. Sur un forum spécialisé, un internaute italien témoigne du marché noir de livrels en Italie et révèle que les fichiers illégaux sont diffusés sous la forme de « *packages* » compressés (au format .rar ou .zip en général) ; ceux-ci contiennent à la fois le livre au format pdf, mais aussi rtf, mobi, lit, html, epub et txt : « *Ces versions ont été scannées, relues et corrigées par au moins trois personnes différentes.* » Les *packages* comportent également des informations sur l'auteur, sa biographie et bibliographie, des liens

⁴⁷ http://en.wikipedia.org/wiki/Ebook_scene.

⁴⁸ A lire le billet de Joël Faucilhon : <http://www.lekti-ecriture.com/bloc-notes/index.php/post/2008/10/04/Portrait-du-pirate-en-conservateur-de-bibliotheque>

⁴⁹ A titre d'exemple, la rubrique « ebooks » de l'un des principaux portails pirates francophones se répartit comme suit (août 2009) : 6330 fichiers de presse et magazines contre 2060 fichiers de livres (soit 75 % contre 25 % du total).

⁵⁰ Ainsi Doopes.com contient une rubrique « ebooks » mais qui concerne à 99 % la presse.

⁵¹ Voir p.19 pour plus de détails sur ces fichiers.

vers des critiques de ses ouvrages, des informations diverses extraites de Wikipedia, et surtout un lien vers une librairie en ligne qui vend le livre au format papier.

A l'inverse, des utilisateurs témoignent de la mauvaise qualité de beaucoup de fichiers pirates :

« Vous ne savez pas ce que vous allez trouver. 90 % des livres scannés requièrent un important travail d'adaptation afin de les rendre compatibles avec les lecteurs d'ebooks (par exemple : des fichiers PDF dont le texte est minuscule ; de multiples tailles et types de polices de caractères (titres de chapitres en taille 50 et corps de texte en taille 8), pas de tabulations, lignes blanches entre des paragraphes, etc.) Au moins 20 % sont complètement illisibles (mauvais scans, formats non convertissables...). Il faut se procurer plusieurs copies d'un même livre pour être sûr de pouvoir en utiliser une. Des ebooks commerciaux, quand ils sont bien formatés et sans DRM, ont donc une valeur ajoutée réelle. »⁵²

Certaines de ces équipes semblent avoir intégré des principes propres au livre, et témoignent des initiatives visant à « libérer le livre » sur Internet. C'est le cas d'une équipe très réputée qui définit son action ainsi :

« Nous sommes une équipe dédiée à la création d'ebooks distribués via les réseaux Peer-to-peer. Sur ce modeste portail, vous pourrez trouver un forum de discussion, une liste à jour de nos distributions mais aussi des informations sur les techniques employées pour créer nos ebooks. Notre mission : Communiquer l'amour de la lecture et de la langue française à tous, au delà de l'éloignement ou du manque de moyens ! Nous ne sommes pas particulièrement portés sur une certaine catégorie de livres, mais faisons plutôt le choix de convertir en ebooks nos coups de cœur. Certains livres copyrightés et récents font partie de nos distributions, comme ce sont souvent ceux qui intéressent le plus grand lectorat. Nous pensons qu'un ebook ne saurait remplacer le livre lui-même, mais simplement donner envie d'aller plus loin avec son auteur(e). Bien entendu, si un ayant droit se pensait réellement lésé par l'existence d'un ebook issu d'un livre de son catalogue, nous cesserons immédiatement de le partager. »

Car de nombreux manifestes et appels au piratage des livres circulent sur le web. C'est le cas de : « Libérez vos livres ! Le guide pour scanner des livres et des magazines »⁵³, traduction d'un texte rédigé originellement en anglais (« Free your books. A guide to scanning books and zines. ») Ce document anonyme, notamment publié sur Scribd, a été lu plus de 24 000 fois.

Une logique assez similaire sur cet autre site qui annonce :

« Vous pouvez utiliser nos ebooks de manière totalement libre, à partir du moment où il s'agit d'une utilisation non commerciale et non professionnelle (lorsqu'il y a la mention "auteur contemporain", le texte n'est pas libre de droit et son utilisation, autre que privée, est soumise à une demande d'autorisation auprès de l'auteur ou l'éditeur). Nous mettons également à votre disposition les sources de nos ebooks, au format .pdf ou .doc, compressé. Nous sommes tous des bénévoles, notre credo est la culture libre et gratuite, mais... nous devons payer l'hébergement du site, les logiciels nécessaires pour élaborer les ebooks, parfois les livres que nous scannons, etc. Et vous avez remarqué l'absence totale de publicité sur le site... Si vous pouvez verser quelques euros chaque année (entre 2 et 5 euros, selon l'état de votre bourse, somme réellement minime par rapport à ce que nous vous offrons), cela nous permettra de continuer de fonctionner. »

⁵² Traduction d'un post sur : <http://www.mobileread.com/>

⁵³ http://www.scribd.com/doc/126099/Liberez-Vos-Livres?_user_id=-1&enable_docview_caching=1

Les motivations des pirates, si elles sont éclectiques, ont pour point commun de vouloir mettre à disposition des ressources non disponibles. Le temps et l'énergie consacrés à la diffusion et au partage de contenus peut dès lors être considérable et, comme l'a justement analysé Alf Rehn⁵⁴, les pirates sont rétribués non pas économiquement mais en valorisation symbolique, par la reconnaissance de la communauté d'internautes à laquelle ils appartiennent.

Dans le domaine du **scantrad*** ou scanslation, c'est-à-dire la traduction et la diffusion illégale de mangas sur le web, la philosophie est souvent assez stricte, comme le montre cette mention trouvée sur un blog spécialisé :

*« Ce blog a été créé pour vous faire découvrir des mangas non sortis en France !
Autrement dit, dès qu'un manga est licencié, il ne sera plus disponible en
téléchargement. » [sic]*

Autrement dit, la plupart des sites de scantrad* placent une certaine limite au téléchargement illégal et ne s'autorisent à diffuser que des ouvrages non disponibles en France. Cette « déontologie » fait qu'ils ont été longtemps tolérés par les éditeurs⁵⁵. De nombreux fans estiment que leurs traductions sont souvent meilleures que les traductions officielles, mais les avis restent très partagés selon les éditeurs et les publications.

Toutefois l'engagement des équipes dans le piratage des livres doit être relativisé. S'il n'est pas contestable que certains sont véritablement dans une démarche de promotion de la culture pour tous, d'autres suivent des stratégies beaucoup moins idéalistes : les fichiers de livrels, et notamment de magazines très grands publics qui n'intéressent qu'une très faible minorité des visiteurs de sites *warez**, sont parfois piratés et diffusés dans le seul but d'améliorer le ratio upload/download. Les règles sont variables, mais un minimum de 1 Ko d'upload pour 3 ou 5 Ko de download est souvent requis afin de pouvoir continuer à faire partie d'une communauté de pirates.

Enfin, il convient de préciser qu'une grande partie des fichiers de livrels ne proviennent pas d'un travail d'équipe mais d'individus isolés, dont les motivations sont assez éclectiques⁵⁶. Le travail des pirates a tendance également à se réduire avec le développement des offres numériques. Une part de plus en plus importante de fichiers numériques circulent (PDF d'imprimeurs, d'éditeurs ou d'auteurs, fichiers électroniques achetés légalement). Sans compter que beaucoup d'internautes mettent à disposition des titres qu'ils n'ont pas scannés eux-mêmes, mais téléchargés sur d'autres réseaux.

⁵⁴ Voir <http://www.alfrehn.com/academic/page2/assets/Rehn%20JSE.pdf>

⁵⁵ Ce n'est pas une règle générale et, devant l'amplification du phénomène, certains éditeurs ont signé une déclaration commune et entrepris des actions contre les plus gros sites de partage, cf. *Livres-Hebdo* du 10/04/2009 : « Une épidémie de scantrad ».

⁵⁶ Une analyse des usages des pirates et des utilisateurs viendra compléter ultérieurement la présente étude.

Processus de piratage d'un livre papier

Si dupliquer illégalement un fichier numérique est un processus extrêmement simple et rapide, pirater un livre depuis sa version papier, comme c'est encore bien souvent le cas, est aujourd'hui une activité complexe qui requiert une forte motivation. En voici les principales étapes⁵⁷.

1) Scan et OCR

Le livre doit être scanné page par page. Les scanners peuvent aujourd'hui être partiellement automatisés afin d'accélérer le processus. Celui-ci demeure toutefois assez long, environ 45 min pour 100 pages. Soit environ 27 sec par page en moyenne en 300 dpi (résolution suffisante pour un livre composés de lignes en noir et blanc ; 600 dpi est préférable pour un ouvrage illustré ou dont la composition est plus riche) avec un scanner de nouvelle génération. Avec un scanner professionnel et un livre non relié, donc permettant un scan entièrement automatique, ce chiffre descend à moins de 3 sec / page soit moins de 5 min pour un livre de 100 pages⁵⁸.

L'OCR est l'acronyme anglais de « optical character recognition » (reconnaissance optique de caractères). Il s'agit d'un procédé informatique qui permet de convertir une image, notamment celle d'une page de livre scannée, au format texte. Le procédé s'est beaucoup amélioré ces dernières années mais il n'est pas encore parfait. Il requiert généralement un fastidieux travail de post-traitement afin de corriger les erreurs et comparer mot à mot le document avec le livre original. Il faut compter environ 1 heure pour 100 pages afin d'obtenir un résultat satisfaisant pour ceux qui accomplissent ce travail de relecture.

2) Traitement de texte et retouchage

Une fois le texte correctement récupéré, il s'agit de le remettre en forme dans un logiciel de traitement de texte ainsi que de le convertir dans le format voulu (en général PDF). Une nouvelle relecture peut également intervenir ici⁵⁹. Cette étape prend de 30 min à 1 heure pour 100 pages.

3) Packaging et upload

Cette étape varie énormément selon le type de piratage. Lorsqu'il s'agit d'une équipe organisée, elle crée un fichier .nfo qui contient les informations essentielles sur sa *release**. Le nom du fichier est formaté selon un standard très précis. Pour se faire reconnaître, les équipes ajoutent leur nom au titre du fichier. Par exemple : *Choc.N108.French.Retail.eBook-SCaN*. Le fichier est ensuite uploadé sur les réseaux peer to peer* et la plupart du temps indexé sur des annuaires ou forums spécialisés.

Au total, on peut estimer jusqu'à **6 h à 10 h** le temps nécessaire pour réaliser l'intégralité d'un piratage d'un livre simple de 200 pages. Ce chiffre peut paraître important, mais il est à comparer au temps passé par les pirates au début des années 2000 pour réaliser des *releases** de vidéo : entre 8 h et 15 h

⁵⁷ Les durées données ici sont approximatives et dépendent de l'état du livre papier, de la qualité des logiciels utilisés ainsi que du degré d'expérience de l'utilisateur.

⁵⁸ Un des nombreux guides en ligne de scan : www.ebooksgratuits.org/guides/tutorial_scanner.htm. Le piratage de la bande dessinée est le secteur le plus en pointe, avec des tutoriaux extrêmement complets et rigoureux.

⁵⁹ Il est à noter qu'un taux d'erreur de 0,01 % dans les corrections correspond déjà à plus de 50 erreurs pour un livre moyen. Ce taux est, selon les usages, d'environ 0,002 % dans l'édition imprimée.

pour un seul fichier. A relativiser toutefois par le fait qu'après quelques réglages initiaux il suffit de laisser l'ordinateur réaliser tout seul l'encodage vidéo. Il n'est pas besoin, contrairement au livre, d'un travail continu. Dans le cas du scantrad*, ces durées sont beaucoup plus importantes et peuvent aller de **25 à 40 h** pour 1 seul tome de manga, en incluant la traduction (10 à 20 h)⁶⁰.

Il faut ainsi considérer un facteur psychologique : plus le fichier est compliqué à pirater et plus, s'il y a une forte demande pour ce fichier, l'opération se transforme en défi dont la réalisation assure reconnaissance et fierté au pirate.

Ces différentes étapes correspondent donc à différentes fonctions au sein d'une équipe, variables selon la taille et l'organisation de celle-ci :

- **La source** : Certaines équipes, parmi les plus importantes, disposent d'un « fournisseur » qui a accès directement aux bandes dessinées avant leur sortie publique. C'est le cas de certains vendeurs en librairie, de collaborateurs de maisons d'édition, etc. En échange d'un accès à d'importantes ressources piratées, la source fournit à l'équipe des fichiers en avant-première. Ce système très courant dans la musique et la vidéo est toutefois moins répandu pour les livres et les BD : d'une part parce que l'œuvre originelle n'est pas numérique et donc difficilement transmissible ; d'autre part parce que la pression de la nouveauté est, à l'exception de quelques best-sellers, moins forte dans l'édition que dans d'autres secteurs.

- **Le scanneur** : De même que l'encodeur s'occupe de convertir le fichier vidéo en un format adaptable au téléchargement sur le web, le scanneur s'occupe de convertir le livre papier en format numérique⁶¹. La minutie et la rigueur de son travail est très variable. Certains scanneurs passent plus de 6 heures pour traiter un épisode d'une bande dessinée, retravailler les couleurs et la résolution. On appelle « rescan » l'opération qui consiste à scanner de nouveau un livre considéré comme mal scanné la première fois. Cette opération peut donner lieu à une nouvelle *release**.

- **Le traducteur** : Le cas échéant, un traducteur peut être nécessaire pour diffuser un livre écrit en langue étrangère. Ce fut le cas d'*Harry Potter* dont la traduction réalisée par des fans et mise à disposition sur le web avant même la sortie de la traduction officielle a fait couler beaucoup d'encre⁶². Ce rôle est crucial dans le domaine du scantrad*⁶³.

- **Le packager** : Le packager conçoit la *release** à proprement parler et génère un fichier brut d'informations (le plus souvent d'extension .nfo ou .txt). Si besoin, il rassemble des fichiers multiples (par exemple plusieurs planches d'une bande dessinée au format image) sous un même fichier compressé. On appelle « repack » l'opération consistant à recomposer un package corrompu ou contenant des erreurs.

- **Le diffuseur (« courrier » en anglais)** : Il est chargé de la diffusion du livre piraté sur le réseau. C'est lui qui charge le fichier sur un serveur et le référence sur les *trackers**. Il contribue à forger la notoriété d'une équipe de *releasers**.

⁶⁰ A noter que ces durées restent des estimations hautes, applicables dans le cas de traitement complet d'un fichier. Il est possible de scanner un livre de 200 pages et de le publier sur Internet en moins de 2 h, au format image.

⁶¹ Dans le cas d'un fichier numérique préexistant, on peut ajouter ici une fonction de « cracking », qui consiste à détourner les éventuels DRM*. Ce cas de figure est encore extrêmement rare pour le livrel.

⁶² En juillet 2005, le tome VI avait été piraté par une équipe de fans et diffusé en moins de 12 heures. Cf : <http://www.01net.com/article/285285.html>

⁶³ A noter que dans le cas du scantrad*, les équipes incluent parfois un « redessinateur » qui, lorsque du texte à traduire chevauche des images, se charge de retravailler les dessins.

```

LWEDDDftFKDLji:                               :ifGDKKEEDDDGLji:
:EDDLi.  :;tfgDKKEDGLfjt;:::.....:;itjFLGDKKKEEDDDGLft;:
DDLi:    .,itfLGGDEEEKKKKKKKKKKKKKEEEEEEDDDDDGLfjt;:
t;       .:;itjffLLLLLLLLLGGLLLLLffjt;:;
          ..:;
          .....

```

Les psy
↓
TOI
↓
Quel est votre problème?

Ouvrage

Scénariste.: Bédou - Cauvin Editeur.: Dupuis

Résumé

Séance de psychanalyse comme si vous y étiez héhé

Releaz

Releazer.: Doc whêshe Nbre de pages.: 46
 Parution.: 20/09/2003 Résolution....: 150ppp
 Format...: JPEG Taille.....: 27.8 Mo

Remerciements

Merci à toute la communauté bdz francophone !
 Les teams, les releazers indépendants ou anonymes,
 Les forums, B██████████, ...
 Et tout particulièrement à :
 ██████████
 ██████████
 les membres de ██████████, ...

Ainsi qu'à toute personne qui investit de son temps
 pour que vive et perdure la communauté bdz !

```

00000 0      000  0 0 0  0  0
   0  0      0      0 00  0  0
00000 00000 00000 0  0 000 00000

```

BDZ Ordre Non Team

```

-----
- Titre . : Fantastic Four -
- ..... : La Renaissance -
- ..... : des Heros -
- ..... : No 1 -
-----
- Editeur .... : Marvel France -
- ..... : MARVEL -
- Année ..... : 1998 -
-----
- 1 ..... : FF Vol II 1 -
- 2 ..... : Cap America -
- ..... : Vol II 1 -
-----
- Langue ..... : Français -
-----
- Scan ..... : ████████ -
- Package .... : ████████ -
-----
- Type ..... : BD -
- Genre ..... : comics -
- Date ..... : 05/03/2005 -
- Format ..... : JPG Q4 Pass 3 -
- Res. .... : 150 DPI -
- Nb de pages. : 100 -
-----

```

Exemples de fichiers .nfo d'équipes de BDZ (bande dessinée pirate) francophones

Méthodologie de l'étude

Pour conduire notre étude, nous avons commencé par identifier les circuits de diffusion des contenus piratés, puis nous avons essayé d'évaluer la quantité de ressources disponibles afin d'en extraire des échantillons analysables.

Les circuits de téléchargement illégal

Nous ne donnerons pas ici une liste de sites offrant un accès à des fichiers illégaux. Mais plutôt un classement des grandes catégories de sources possibles. Chacune de ces catégories a ses propres logiques de recherche, et donne un accès privilégié à des types de fichiers particuliers.

- 1) ed2k : le P2P* via le réseau eDonkey*
- 2) Torrent : le P2P* via le réseau torrent*
- 3) DDL : le Direct Download* ou téléchargement direct
- 4) Usenet : les newsgroups ou groupes de discussion
- 5) IRC : l'Internet Relay Chat
- 6) HTTP : les recherches sur le web

Ces différentes sources correspondent à des usages d'Internet très différents.

Utilisateur novice ou occasionnel	HTTP, DDL*
Utilisateur averti	DDL*, eDonkey*, BitTorrent
Utilisateur expérimenté	Usenet, IRC, FTP, <i>trackers*</i> privés

D'une façon générale, et à l'exception de quelques best-sellers très diffusés, il reste assez fastidieux de trouver sur internet la version pirate d'un livre précis. Mais l'amélioration de l'accessibilité semble être une tendance de fond depuis un an ou deux.

1) Ed2k : le piratage via le réseau eDonkey*

eDonkey* est un logiciel de partage de fichiers en peer to peer*, c'est-à-dire de machine à machine. Les sites dits de *dumplinks* ou *indexers* proposent des liens de téléchargement sous la forme d'annuaire thématique : musique, films, séries, jeux vidéo... Des annuaires de plus en plus nombreux comportent une rubrique « ebooks », qui indique les fichiers rendus disponibles gratuitement par les utilisateurs.

Nous avons utilisé le logiciel eMule, l'interface eDonkey* la plus utilisée dans le monde. Le logiciel inclut une fonction de recherche qui permet l'utilisation de critères croisés.

Le protocole de recherche utilisé sur eMule est très important car il permet de filtrer les résultats avec pertinence. Ainsi une grande partie des fichiers comporte un **mot-clé générique** qui permet de préciser s'il s'agit d'un livrel, d'une BD ou d'un livre audio.

- > pour les ebooks : ebook[s] french, ebook[s] francais, ebook[s] FR, ebook[s] livre[s], livre[s]
- > pour les livres audio : livre audio, audio[]book[s] French, audio[]book[s] FR
- > pour la bande dessinée : bande dessinee, BD french, BD francais, comics French, BDFR

L'utilisation de cette méthode introduit nécessairement un biais du fait que l'internaute n'effectue que rarement une recherche sur un mot-clé générique, mais plutôt sur le titre d'un livre ou un nom d'auteur. Toutefois, elle reste pertinente car l'on constate que la majorité des fichiers sont « taggués » (identifiés par un tel mot-clé) de cette manière, ne serait-ce qu'au moins une des versions pirates d'un même livre, lorsqu'il en existe plusieurs. Afin d'obtenir un échantillon de taille raisonnable à analyser, nous avons effectué ces recherches sur les principaux serveurs ed2k*, en supprimant les fichiers ne disposant d'aucune source.

A ces mots-clés ont été associés les extensions de formats de fichiers les plus répandus : .pdf (format PDF), .doc (format Word), .txt (format texte brut), les fichiers d'images .jpg et les fichiers compressés .rar (WinRar) et .zip (WinZip).

Il est enfin possible d'affiner la recherche en utilisant les noms d'éditeurs ou les titres des ouvrages, mais par définition ce mode de recherche est quasi illimité. Ce serait l'objet d'une étude complémentaire : aller chercher sur ces réseaux uniquement un corpus précis de livres choisis en fonction de critères eux aussi définis précisément. Comme nous le verrons plus loin, nous avons néanmoins pris en compte ce type de démarche et recherché les convergences entre l'offre illégale et l'offre légale.

2) BitTorrent : le P2P* via le réseau torrent*

Torrent* est également un protocole d'échanges de fichiers en peer to peer*, dont le succès a un peu rattrapé celui d'eDonkey*. Le fonctionnement est toutefois très différent car il suppose, pour un pirate qui souhaite diffuser une *release**, de créer un fichier torrent* (*tracker**) et de le faire référencer sur un annuaire spécialisé. Cette création et ce référencement, s'ils sont simples, rendent tout de même la procédure plus compliquée que via eDonkey.

Pour effectuer une recherche sur les torrent*, il faut donc accéder à des annuaires ou agrégateurs qui répertorient les liens et les classent en différentes catégories thématiques, dont les ebooks. Ces portails disposent parfois de statistiques de téléchargements.

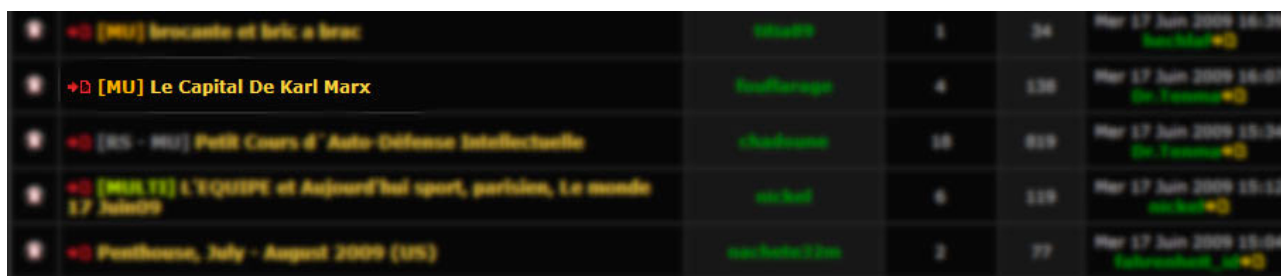
La méthode d'identification des fichiers utilisée est donc l'extraction exhaustive des fichiers inclus dans ces catégories « ebooks » des principaux agrégateurs.

3) DDL : le *direct download* ou téléchargement direct

Le *direct download*⁶⁴ est en progression exponentielle depuis deux ou trois ans, car il offre à la fois une alternative sécurisée face à la surveillance du piratage, et des vitesses de téléchargement souvent plus élevés.

L'apparition de sites de stockage comme MegaUpload, RapidShare ou le service d'hébergement de Free, qui proposent des espaces de plusieurs gigas pour y déposer des fichiers dont la légalité n'est que rarement vérifiée, a contribué au développement de ce mode de téléchargement. Des sites spécialisés, en général sous la forme de forums nécessitant une inscription, fournissent des listes renouvelées quotidiennement de liens de téléchargement direct. Ces liens sont généralement standardisés : le titre commence par le nom de l'hébergeur entre crochets, suivi par le titre du livre proposé.

Exemple : [MU] *Le capital de Karl Marx* : il s'agit d'un lien de téléchargement du livre *Le Capital*, hébergé sur MegaUpload. [MULTI] signifie que les liens de téléchargement sont disponibles sur plusieurs sites de stockage différents.



Titre	Statut	Nombre de téléchargements	Nombre de téléchargements	Date
[MU] <i>Le Capital de Karl Marx</i>	upload	4	136	Mer 17 Juin 2009 14:47
[MU] <i>Le Capital de Karl Marx</i>	upload	4	136	Mer 17 Juin 2009 14:47
[MU] <i>Le Capital de Karl Marx</i>	upload	4	136	Mer 17 Juin 2009 14:47
[MULTI] <i>Le Capital de Karl Marx</i>	upload	4	136	Mer 17 Juin 2009 14:47
[MULTI] <i>Le Capital de Karl Marx</i>	upload	4	136	Mer 17 Juin 2009 14:47

Capture d'écran d'un forum de warez

La méthode d'identification des fichiers est la même que pour BitTorrent : l'extraction des fichiers contenus dans les rubriques « ebooks » de principaux sites et forums de référencement.

Des moteurs de recherche spécifiques aux hébergeurs de DDL* ont vu le jour, comme www.filecrop.com ou encore www.filestube.com. Toutefois, ces moteurs ne retournent que les résultats contenant le mot-clé de la recherche dans le titre du fichier, ce qui est de fait très limitatif. Par ailleurs, les fichiers sont filtrés très régulièrement et les contenus illégaux supprimés, ce qui augmente considérablement le taux de liens de téléchargement référencés mais non accessibles.

Il existe un autre moyen plus surprenant de trouver des fichiers sur le web. Il consiste à utiliser les fonctionnalités de recherche avancée de Google. On appelle parfois ces dernières « Google hacks » car elles peuvent, dans certains cas, permettre l'accès à des données censées être non publiques. Par exemple, la ligne suivante indexe des listes de sites proposant des livres en téléchargement :

```
intitle:"index.of" ebooks, intitle:"index of" +("/ebooks"|"book") +(chm|pdf|zip)
```

Cette dernière méthode reste anecdotique et nous n'en avons pas tenu compte dans nos recherches de fichiers.

⁶⁴ Le DLL est en réalité une forme de téléchargements en HTTP puisqu'il s'agit de liens directement accessibles depuis un navigateur web. Toutefois ses spécificités justifient une catégorie indépendante.

4) Usenet : les newsgroups ou groupes de discussion

Plus connu en France sous le nom de « newsgroups », il s'agit d'un réseau de forums de discussion, existant depuis 1979, et dont Google Groups n'est que la partie grand public la plus visible aujourd'hui. A l'origine développés pour diffuser du texte brut, les newsgroups sont réputés pour leur capacité à véhiculer des fichiers binaires*, par exemple des fichiers de musique, des films ou des livrels.

Les noms des groupes de téléchargement de livrels francophones contiennent généralement le mot « binaries » (fichiers binaires*) ainsi que le mot « ebook » et « french ». Des moteurs de recherche spécifiques, sur le web, comme www.binsearch.info, permettent de trouver plus facilement les fichiers diffusés sur Usenet.

La méthode d'extraction a été identique à celle du réseau Ed2k, c'est-à-dire l'utilisation de ces mots-clés génériques.

5) IRC : l'Internet Relay Chat

L'Internet Relay Chat ou IRC est un protocole de communication par texte instantané.

Le téléchargement par IRC est réservé à une minorité d'utilisateurs chevronnés, car il requiert des compétences informatiques dépassant largement la moyenne de celles de la population. Il suppose la connexion à un canal (« channel »), nécessitant souvent une inscription préalable⁶⁵. Il s'agit ensuite d'opérer un certain nombre de commandes pour la recherche et le téléchargement. Devant l'impossibilité de trouver des « annuaires » d'ebooks illégaux sur IRC, la méthode utilisée a été la même que pour ed2k et Usenet. Plus précisément, on se reportera aux commandes de recherche comme @find ou @search suivies du titre du livre et aux commandes de téléchargement sous la forme : !serveur Bernard Werber - Les Fourmis [Html].rar ::INFO:: 167.0KB)

6) HTTP : les recherches sur le web

Selon Google, Google Books comporterait environ 10 millions d'ouvrages, dont 2 millions du domaine public, 2 millions publiés avec l'autorisation des ayants-droit, et 6 millions épuisés en version imprimée mais dont les droits ne sont pas encore négociés.

Ayant fait largement parler de lui auprès du grand public quand Barack Obama l'a utilisé pour diffuser des documents pour sa campagne électorale, le site **Scribd** (www.scrib.com), créé en 2007 aux États-Unis, permet de publier ses textes dans une base de données accessible à tous. Le site supprime systématiquement tout document illicite sur simple demande des ayants droit, mais il reste encore une poignée d'ouvrages francophones disponibles au visionnage et au téléchargement en toute illégalité.

Les mêmes remarques peuvent être faites sur **WattPad** (www.wattpad.com), un logiciel de lecture de livrels sur téléphones portable lancé en 2006, lui aussi très critiqué pour héberger sans accord préalable des fichiers sous droits.

Il existe en outre des moyens alternatifs pour trouver des livrels. Par exemple, une simple recherche sur YouTube (www.youtube.com) sur les mots-clés « free ebooks » donnera des résultats surprenants,

⁶⁵ Les *channels* anglophones les plus populaires portent les noms évocateurs de #ebooks, #bookz et #audiobooks.

comme des vidéos proposant des méthodes, plus ou moins efficaces, pour chercher des livres sur le web⁶⁶.

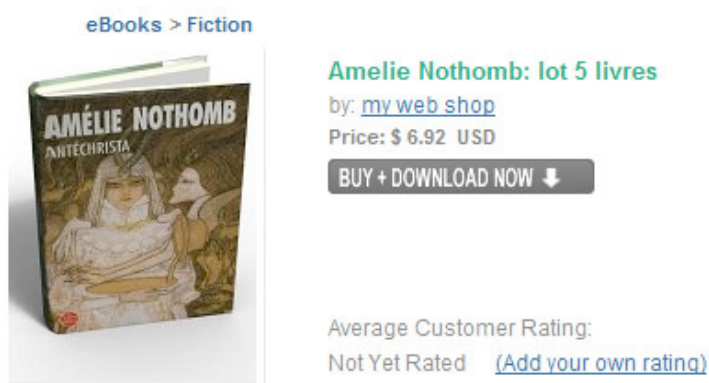
Une autre méthode intuitive consiste à entrer directement les mots-clés recherchés dans Google. Et les résultats sont parfois étonnants. Le tableau suivant indique le nombre de sites proposant des téléchargements illégaux de livres de quelques auteurs de best-sellers, sur une recherche « Nom de l'auteur + ebook » sur les 10 premiers résultats :

Bernard Werber	6 / 10
Amélie Nothomb	2 / 10
Marc Levy	2 / 10

Recherche sur Google.fr au 18/06/2009.

Alternative plus simple aux requêtes brutes sur Google, des moteurs de recherche spécialisés comme PDF Geni ou PDF Search Engine (moins performant et moins ergonomique que le premier), permettent d'effectuer des recherches directement sur des fichiers au format PDF. Les résultats n'aboutissent pas tous, loin s'en faut, à des fichiers illégaux, mais ces outils indexent les contenus sans distinction.

Plus étonnant, et très rare, certains sites proposent des livres payants, au format PDF. La capture d'écran est celle d'un site de partage et de vente de fichiers, qui fonctionne par paiements Paypal, et propose par exemple à son catalogue un lot de 5 livres d'Amélie Nothomb en numérique pour 6,92 \$.



The screenshot shows a product listing for 'Amélie Nothomb: lot 5 livres' on a website. The page header includes 'eBooks > Fiction'. The product image shows a book cover for 'AMÉLIE NOTHOMB ANTECHRISTA'. The price is listed as '\$ 6.92 USD'. There is a 'BUY + DOWNLOAD NOW' button with a downward arrow. Below the product, it says 'Average Customer Rating: Not Yet Rated' with a link to '(Add your own rating)'.

Ce fut également le cas du site d'enchères eBay qui a longtemps proposé à la vente des livrels sous droits. Mais il a depuis supprimé la très grande majorité de ces enchères illégales. En témoigne un récent rapport de l'AIDRO, l'organisme de gestion des droits des auteurs en Italie⁶⁷ (« Relazione sull'attività antipirateria AIDRO 2008 »), qui a fait supprimer plus de 850 offres de livres piratés.

⁶⁶ http://www.youtube.com/results?search_type=&search_query=free+ebook&aq=f

⁶⁷ <http://www.aidro.org/Portals/26/File/Documenti/Relazione%20antipirateria%20AIDRO%202008.pdf>

Comment évaluer le nombre d'ouvrages disponibles illégalement et de téléchargements ?

Chaque circuit de diffusion correspond à une logique de comptage de fichiers relativement spécifique. Notre étude portant sur l'offre et non la demande, nous n'avons considéré ici l'évaluation du nombre de téléchargements qu'à titre purement indicatif.

A) Ed2k*

Le réseau ed2K n'offre pas de statistiques pertinentes. A chaque ouvrage est associé un nombre de « sources », qui correspond aux « clients » (utilisateurs) partageant le fichier à un moment donné (pour les livrels, ce nombre est généralement très faible, entre 0 et 1). A noter que ce nombre n'offre aucune indication quant au nombre de téléchargements complets du fichier correspondant. On ne peut donc proposer que des classements par nombre de sources (cf. annexe 5).

B) BitTorrent*

Il est envisageable de relever et d'agrèger certaines données des sites de *trackers*⁶⁸, même si ces chiffres sont à prendre avec précaution. En ce qui concerne les téléchargements, il est possible d'extrapoler grâce à l'audience supposée des sites considérés. C'est ainsi qu'en prenant les 3 ou 4 grands agrégateurs de liens torrent, comme Mininova, The Pirate Bay ou Demonoid⁶⁹, il est possible d'obtenir des chiffres de téléchargements relativement pertinents. C'est en général ces chiffres qui sont relayés dans les études⁷⁰. Mais vu la faible part des torrent* dans le téléchargement des livrels francophones, ces calculs restent secondaires. Ce n'est pas le cas par exemple dans le secteur de la vidéo où le protocole BitTorrent recouvre encore la très grande majorité des téléchargements⁷¹.

C) Direct download*

Là encore, il est impossible d'obtenir des données exhaustives. La seule opération cohérente consiste à agréger les données des principaux portails référençant les liens de téléchargement direct. On a pu estimer ces portails entre 10 et 20 pour le livrel francophone, chacun proposant entre 100 et 2 000 liens de téléchargement. Certains proposent des statistiques de téléchargement, mais celles-ci ne sont pas toujours fiables.

Par ailleurs, dans le livre encore plus que dans la musique et certainement la vidéo, le nombre de téléchargements ne signifie pas grand chose. La faible taille des fichiers ainsi que la forte présence de « packages », c'est-à-dire de plusieurs livrels à télécharger ensemble (cela peut parfois aller jusqu'à plusieurs milliers de livres), facilite le téléchargement de nombreux livres qui ne seront toutefois pas tous consultés. Beaucoup seront seulement stockés et détruits au bout d'un certain temps de non-utilisation.

⁶⁸ C'est le cas de Demonoid (www.demonoid.com) qui associe à chaque fichier le nombre de *seeders**, de *leechers**, mais aussi de téléchargements complets.

⁶⁹ Classement des principaux agrégateurs (rang des sites les plus visités sur le web selon Alexa.com) : Mininova (94^e), The Pirate Bay (110^e), TorrentReactor (360^e), btjunkie (380^e), Demonoid (596^e).

⁷⁰ Par exemple : <http://freakbits.com/the-10-most-pirated-ebooks-of-2009-0831>. Cette étude révèle que le top 3 des livrels anglophones les plus téléchargés en torrent* sont : 1) *Kamasutra*, 2) *Adobe Photoshop Secrets*, 3) *The Complete Idiot's Guide to Amazing Sex*.

⁷¹ Ce qui permet à un organisme de veille comme BigChampagne (<http://bcdash.bigchampagne.com>) de publier les chiffres suivants : <http://www.thresq.com/2009/08/illegal-downloaders-looking-for-heroes.html> (cf. p.44).

La plupart des gros éditeurs font aujourd’hui appel à des cabinets de conseil ou des sociétés spécialisées dans la surveillance des réseaux, qui ont développé des outils spécifiques, à l’instar d’Attributor⁷² qui a signé un contrat avec Hachette en juin 2009 afin de détecter les copies pirates. Si ces sociétés sont efficaces pour détecter le piratage de fichiers spécifiques, les avis sont réservés sur leur capacité à fournir une image globale satisfaisante de l’offre et du téléchargement illégal. La plupart des hackers considèrent que les estimations officielles sont souvent erronées, voire fantaisistes, car les outils de calculs non adaptés et les intervenants techniques pas assez au fait des technologies propres au partage de fichiers.

Quelle que soit la méthode utilisée, les chiffres trouvés ne sont nécessairement que partiels et ne peuvent prétendre à recouvrir l’ensemble de l’offre illégale. Il existe de plus en plus de réseaux privés, sous la forme de FTP ou de « darknet* » (dont le programme le plus connu est Freenet⁷³, un réseau de partage dont l’accès peut être restreint aux seuls utilisateurs connus, ce qui leur garantit l’anonymat. Freenet est d’ailleurs très utilisé pour le partage de textes). Il est impossible d’évaluer correctement le nombre de livrels qui y circulent. Par ailleurs, depuis le procès et la revente de l’agrégateur de torrents* The Pirate Bay, de nombreux sites de peer to peer* craignant d’être fermés à leur tour par les autorités se sont transformés en *trackers** privés, c’est-à-dire uniquement accessibles sur inscription, voire sur invitation. Et dont l’étendue du contenu est par là- même extrêmement difficile à évaluer.

Voici quelques statistiques, remontant déjà à 2007, sur les fichiers de livrels anglophones référencés par les principaux sites de torrents*⁷⁴ :

Trackers BitTorrent	Livrels	Livres audio
www.fulltdls.com	27 900	264
www.mininova.org	15 500	4 000
thepiratebay.org	14 500	559
btjunkie.org	8 000	1 480
www.torrentportal.com	6 200	1 100
www.isohunt.com	4 700	990
www.torrentreactor.net	3 400	926

Inclus tous les fichiers, y compris ceux comportant 0 seeds.
Source : FileShareFreak.com, décembre 2007.*

Mais ces chiffres sont de simples agrégats difficilement exploitables, notamment parce qu’ils référencent aussi des livres qui ne sont plus disponibles au téléchargement. Ainsi, analysons les données fournies par le *tracker** le plus connu, Mininova :

Catégorie	Nombre de torrents*
Articles	2 879
Livres audio	13 468
Comics	7 047
Livres	44 875
Magazines	2 766

⁷² <http://www.attributor.com>

⁷³ www.freenetprojet.org

⁷⁴ D’après le blog : <http://filesharefreak.com/2008/04/08/specialized-private-bittorrent-websites-elearning-ebooks/#more-151>

Manga	212
Manuels	1 670
Jeux de rôles	29

Nombres de fichiers disponibles sur Mininova, 08/07/2009.

La rubrique « Livres » est en réalité une somme de fichiers très composites, ne répondant que partiellement à la définition du livre. Ainsi, sur les 100 résultats les plus récents, on constate notamment que :

- 59 % sont soit des magazines classés dans la mauvaise catégorie ; soit de simples extraits de quelques dizaines de pages ; soit des ouvrages amateurs et non commercialisés (cours d'informatique ou de cuisine, textes de propagande religieuse, etc.) ;
- 28 % n'ont aucun *leechers** (c'est-à-dire ne sont plus disponibles de fait en téléchargement) ;
- 80 % sont des ouvrages (y compris de nombreux modes d'emploi) informatiques et techniques.

Une recherche sur le mot-clé en français « livre » dans la rubrique « ebook » de Mininova donne les statistiques suivantes, sur les 62 résultats obtenus :

Littérature	12	19 %
Essais	3	5 %
Bandes dessinées	3	5 %
Livres audio	12	19 %
Jeux de rôles	2	3 %
Pratique (dont cuisine)	18 (10)	29 % (16 %)
Ouvrages du domaine public	9	14 %
Collection d'ouvrages (fichier contenant plusieurs livrels, par exemple tous les tomes d'Harry Potter)	8	13 %
Ouvrages amateurs ou non identifiés	7	11 %

Source : Mininova, 8 juillet 2009.

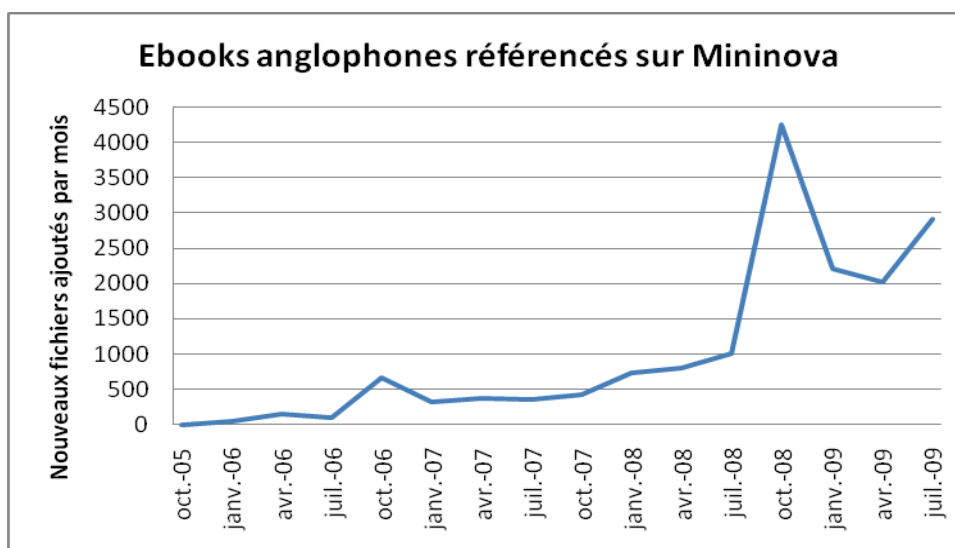
On trouvera également en annexe 4 et 5 les tableaux détaillés des résultats trouvés dans la rubrique « ebooks » en langue française du *tracker** Demonoid ainsi que par une recherche sur le logiciel de peer to peer* eMule.

Evolutions récentes

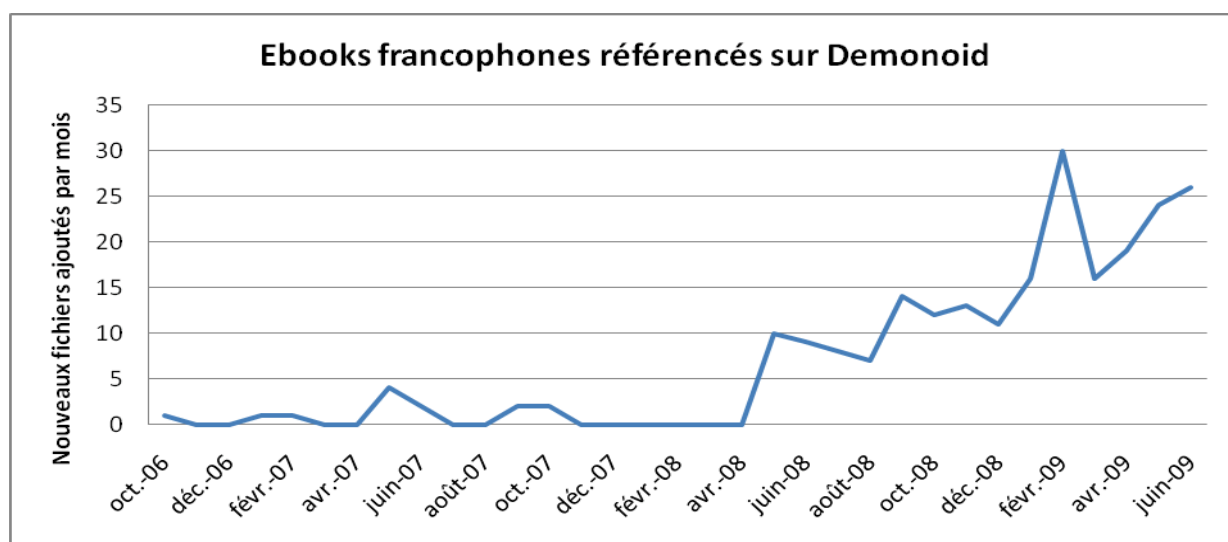
Selon le cabinet d'études britannique Envisional⁷⁵, le nombre de livres piratés s'élevait à environ 7 200 en 2001, entre 25 000 et 30 000 en 2004. Malheureusement, aucune information précise n'a été donnée quant au mode de calcul employé pour parvenir à ces résultats. Par ailleurs, aucune mise à jour n'a été publiée depuis et les chiffres plus récents dont ils disposent sont réservés à quelques-uns de leurs clients.

Il est certain que, depuis moins de deux ans, l'offre illégale de livrels a augmenté fortement, comme en témoignent les deux diagrammes suivants :

⁷⁵ <http://www.envisional.com/piracy>.



Source : Mininova.org.

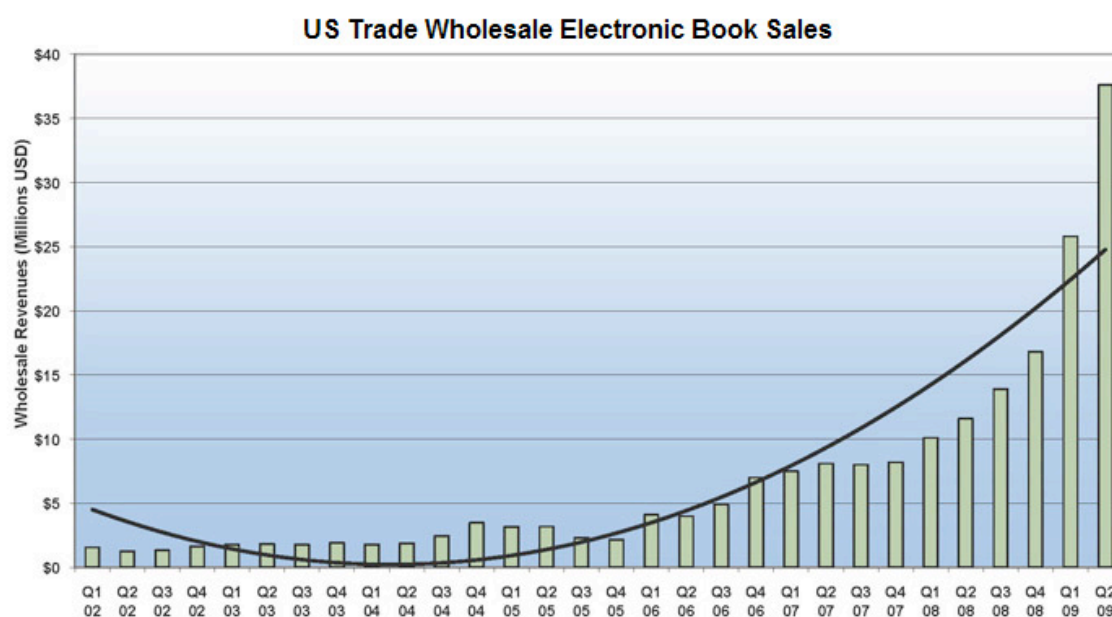


Source : Demonoid.com.

Les deux diagrammes ci-dessus sont difficiles à interpréter car ils reposent sur un faisceau complexe de facteurs :

- Les usages spécifiques des agrégateurs considérés⁷⁶.
- L'évolution des nouvelles technologies et de leurs usages, notamment la progression du peer to peer* avec le développement des connexions haut débit ;
- Des effets de seuil provoqués par certains phénomènes endogènes au marché du livre numérique : forte progression des ventes de liseuses au cours de l'année 2008, lancement de sites de vente numérique (par exemple Relay.com, qui a ouvert une offre de livrels en octobre 2008). Ainsi ces courbes sont comparables à celle des ventes de livrels aux États-Unis :

⁷⁶ Par exemple, le nombre de nouveaux fichiers sur la deuxième courbe (Demonoid) est à 0 de novembre 2007 à avril 2008, période pendant laquelle le site a été fermé par la Canadian Recording Industry Association.



Ventes aux distributeurs/diffuseurs, du 1^{er} trimestre 2002 (Q1 02) au 2^e trimestre 2009 (Q2 09). Source : IDPF.

Sur l'un des plus importants forums de *warez** francophone, la rubrique « ebooks » dépasse même les rubriques logiciels et jeux vidéo en nombre de sujets postés, qui correspondent peu ou prou au nombre de fichiers mis à disposition, car chaque nouvel article donne accès à au moins un nouveau lien de téléchargement. Toutefois, cette rubrique « ebooks » est hypertrophiée par le fait qu'elle indexe également la presse. Elle traduit une certaine variété de l'offre illégale, mais ne reflète pas la réalité en termes de demande : films, musiques et logiciels demeurent à des volumes de téléchargement illégal sans comparaison avec ceux des livrels.

	Nombre de sujets postés	% du total
	76 356	100 %
Musique	9 141	12 %
Ebooks	5 523	7,2 %
Logiciels	4 057	5,3 %
Jeux vidéo	4 033	5,2 %

Statistiques de l'offre illégale au 19/06/09, recueillies sur l'un des plus importants forums de *warez** francophone

Estimation globale de l'offre illégale de fichiers

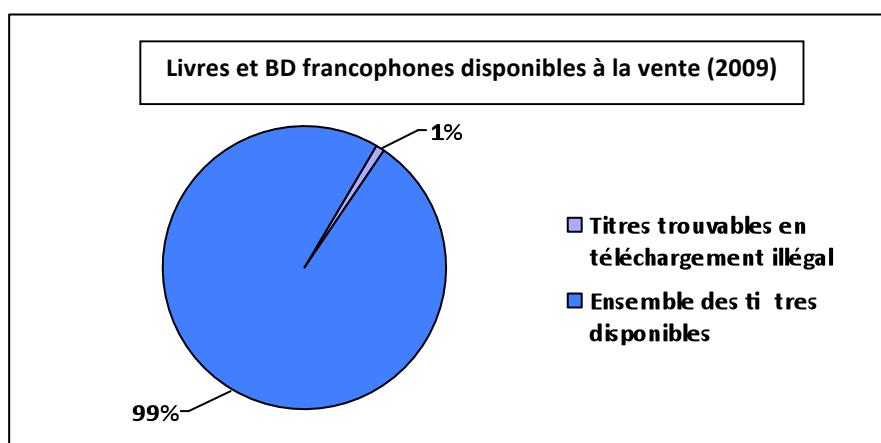
Comme nous venons de le voir, l'estimation de l'offre globale de fichiers fait l'objet de spéculations plus ou moins hasardeuses. Certains acteurs du secteur du livre en France sont persuadés que cette offre est quasi inexistante ; d'autres à l'inverse pensent que la majorité des contenus de l'édition sont déjà piratés.

Après trois mois de recherches sur un très grand nombre de fichiers illégaux à disposition des internautes dans les différents circuits de diffusion, nous donnons les estimations suivantes concernant les titres français :

- ❖ Environ **1 000 à 1 500 titres** de livrels : une grande partie de ces livres sont des livrels scientifiques, techniques ou médicaux ;
- ❖ Environ **3 000 à 4 500 titres** de bandes dessinées : le nombre considérable de titres provient ici notamment de la diffusion de séries complètes (par exemple : l'intégrale de *Lucky Luke*) ;
- ❖ Environ **200 à 300 titres** de livres audio, dont la moitié au moins concerne des textes du domaine public.

Soit un total d'environ **4 000 à 6 000 titres en cumulant livres et BD**, ce qui est très faible au regard du nombre de titres papier disponibles⁷⁷ (sont exclus les ouvrages disponibles seulement sur des réseaux privés, accessibles par une connexion sécurisée, et dont il est impossible d'évaluer la quantité.)

	Titres piratés	Titres disponibles en offre légale papier	% des titres piratés
Livres + BD	4 000-6 000	565 000	0,7 % - 1,1 %



La majorité des fichiers disponibles ne sont que difficilement accessibles, car peu partagés par les internautes (entre 0 et 1 source en peer to peer*). Le téléchargement illégal de livres est un piratage résiduel dans un secteur qui n'a pas encore connu la numérisation. Ce taux de piratage progressera certainement avec la mise à disposition de livrels au format numérique, nécessitant une mise à jour régulière de cette étude.

Enfin un autre phénomène difficilement mesurable est le prêt ou don électronique. Il concerne le partage de livrels qui ne passe ni par les réseaux peer to peer*, ni par le web d'une façon générale, et qui est sans doute à ce jour très important en volume : collègues qui s'échangent, par courrier électronique, des études au sein d'une entreprise ; universitaires qui partagent un ouvrage avec leurs étudiants ; amis qui s'envoient des livres après en avoir parlé sur un réseau social... Ce type de don s'apparente finalement (du moins quant au principe) au prêt ou au don de livres imprimés.

⁷⁷ Pour Jean Arcache, PDG de Place des Editeurs : « Nous avons de la chance par rapport au secteur des jeux vidéo qui est le plus touché. Presque toutes les 50 principales références, qui correspondent à 80 % des ventes, sont piratées. C'est la même chose pour les DVD. En ce qui concerne les CD, 500 références font 80 % des ventes, et là aussi presque la totalité est piratée. Dans le monde du livre, 3 000 références ne correspondent qu'à 50 % des ventes. Or, avant même que ces 3 000 ouvrages soient piratés, on aura le temps de se retourner.» http://www.challenges.fr/actualites/medias/20090820.CHA6503/le_livre_ne_craint_pas_le_net.html

Echantillonnage

A partir de ces observations, nous avons pu établir des ensembles ou échantillons⁷⁸ que nous avons analysés en détail afin de comprendre quelle était la nature exacte de l'offre illégale. Nous avons retenu le nombre de fichiers suivants :

	Nombre de fichiers
Livrels	500
BD	300
Livres audio	100

Ces échantillons ont été récupérés selon une logique d'entonnoir : nous avons commencé par relever les fichiers disponibles facilement (cf. la classification des types de téléchargeurs p.25), c'est-à-dire sur le réseau eDonkey et en BitTorrent. Puis nous avons affiné en utilisant les principaux annuaires de direct download*, ainsi que les sources plus alternatives comme les newsgroups et l'IRC.

Nous nous sommes donc mis à la place d'un internaute moyennement expérimenté, afin de voir ce qu'il était possible de trouver sans des prérequis de savoir informatique trop importants.

La taille des ensembles a été déterminée de façon à conserver le meilleur rapport possible entre une nécessaire représentativité et la complexité d'analyser un grand nombre de fichiers. Au total, les quantités de livres piratés sont suffisamment faibles pour que les ensembles considérés soient assez représentatifs : ainsi, pour les livres, l'échantillon représente environ 1/3 de l'estimation du volume global d'ouvrages piratés disponibles sur Internet.

Les échantillons ont tendance à sous-représenter certaines catégories de livres, trouvables en grandes quantités sur des sites spécialisés, voire privés, mais qui n'apparaissent que dans une moindre mesure lors de recherches généralistes. C'est le cas des ouvrages scolaires, mais surtout des ouvrages de STM*, d'informatique, et de jeux de rôles⁷⁹.

Au final, la répartition par circuits des corpus de fichiers que nous avons étudiés est la suivante :

	Fichiers trouvés	Fichiers retenus	Répartition
Ed2K*	980	522	58 %
Direct download*	578	297	33 %
Torrent*	132	54	6 %
Usenet	77	18	2 %
Autre (IRC, HTTP)	22	9	1 %

Le tableau se lit ainsi : « 58 % des fichiers analysés proviennent du réseau ed2K ». La différence entre fichiers trouvés et fichiers retenus s'explique par l'exclusion de tous ceux qui ne correspondent pas à nos critères, c'est-à-dire essentiellement ne répondent pas à la définition du livre, sont corrompus ou sont des *fakes**

⁷⁸ Les fichiers ont été recueillis pendant une période de trois mois allant du 1^{er} mai 2009 au 1^{er} août 2009.

⁷⁹ Aux États-Unis, un site comme TextBookTorrents (créé en janvier 2007, aujourd'hui fermé) et partageant des milliers de manuels scolaires, a fait couler beaucoup d'encre. On trouve en France des sites équivalents, quoique moins bien fournis.

Si l'on affine le tableau précédent par catégories :

	Ed2K	DDL	Torrent	Usenet	IRC
Total	58 %	33 %	6 %	2 %	1 %
Livres	51,6 %	40,1 %	6,7 %	0,2 %	1,4 %
BD	70 %	20,4 %	2,8 %	6,8 %	0 %
Livres audio	78,3 %	0,2 %	21,5 %	0 %	0 %

L'échange de fichiers en peer to peer sur le réseau eDonkey reste donc encore le moyen le plus utilisé pour échanger illégalement des livres. Mais il est vite rattrapé par le direct download*. Les autres modes de téléchargement apparaissent comme plus marginaux.

Il est à noter que les chiffres ci-dessus ne reflètent pas la diffusion de l'ensemble des textes piratés, mais de ceux qui sont accessibles le plus facilement, selon la méthode utilisée pour créer les échantillons de l'étude. Ils ont donc tendance à sous-valoriser certains circuits moins accessibles aux profanes, à l'instar des newsgroups (essentiellement pour les bandes dessinées).

L'accès à des fichiers par HTTP n'apparaît pas pour deux raisons : d'une part, l'essentiel des fichiers illégaux disponibles sur des sites de partages de publication (comme Scribd) ont été supprimés ces derniers mois. D'autre part, la très grande quantité de mangas disponibles en lecture directe sur le web (dans le domaine du scantrad* essentiellement) font l'objet d'un cas particulier et ne sont pas pris en compte ici, mais dans l'analyse thématique (p.48).

Résultats

Il est nécessaire de rappeler que les résultats de notre étude doivent être pris avec précaution, compte tenu de la volatilité des échanges sur Internet et de la difficulté d'établir des échantillons représentatifs sur des données illégales.

Analyse des fichiers

Qualité

Pour juger de la qualité d'un fichier, une dizaine de caractéristiques ont donné lieu à une évaluation globale. Les caractères du texte sont-ils flous ? Le texte est-il imprimable ? Respecte-t-il la mise en page originale ? Les photos ou illustrations ont-elles beaucoup perdu à la compression ? Y a-t-il des pages ou des fragments de page manquants ? Les pages sont-elles alignées correctement ? Le fichier est-il bien compressé ?

L'évaluation est un chiffre de **1 à 5** :

- 1 : fichier le plus bas de gamme, en général un simple texte au format brut, sans mise en page, difficilement lisible ou incomplet.
- 2 : fichier au format traitement de texte (Word ou Open Office), très légèrement remis en page, ou PDF de très mauvaise qualité.
- 3 : fichier moyen, en général un PDF au format image dont la résolution ne permet pas l'impression.
- 4 : fichier de très bonne qualité, auquel il manque seulement un ou deux critères de confort de lecture (par exemple : recherche en mode texte dans le cas d'un essai).
- 5 : un livre scanné et transformé au format PDF (image ou en texte libre) d'excellente qualité, se rapprochant au maximum de l'original et offrant un confort de lecture optimal⁸⁰.

Remarque : contrairement à la musique et à la vidéo, le taux de *fakes** pour les livrels (sur notre échantillon en tout cas) est extrêmement bas, moins de 1 %. Cela est notamment dû au fait que la pollution du peer to peer* ne fait pas (encore ?) partie de la stratégie anti-piratage des producteurs de contenus légaux.

	1	2	3	4	5	Moyenne
Total	4 %	5,6 %	13,4 %	36,6 %	40,4 %	4
Livres	5,5 %	8,4 %	18,3 %	32,4 %	35,4 %	3,8
BD	1,7 %	0,5 %	4,8 %	50 %	43 %	4,3
Livres audio	0 %	0 %	3,3 %	20 %	76,7 %	4,7

Notre étude montre que la **qualité globale moyenne des fichiers est de 4 sur 5**. 77 % des fichiers

⁸⁰ Il s'agit souvent de fichiers provenant directement de versions numériques « crackées » ; c'est le cas avec beaucoup d'ouvrages de STM publiés chez des éditeurs comme Eyrolles et EDP Sciences.

(67,8 % des livres et 93 % des BD) ont une qualité supérieure ou égale à 4. Ce qui va à l'encontre d'une idée reçue selon laquelle les fichiers illégaux sont de mauvaise qualité.

Formats

Explication des formats :

- **PDF image/texte** : fichiers au format PDF d'Adobe, avec ou sans la possibilité de lire le contenu en mode texte (donc de copier/coller des passages). L'essentiel des ces fichiers proviennent de scans réalisés manuellement depuis des ouvrages papier. Une faible quantité (que l'on peut évaluer à moins de 4 %⁸¹) provient du piratage de fichiers numériques préexistants disponibles en offre légale, et dont il est quasi impossible de définir à quel niveau s'est effectuée la fuite (acheteur, service de presse de l'éditeur, correcteur, auteur, etc.)
- **Word** : fichiers au format .doc ou .rtf (dans quelques très rares exceptions, au format .odt d'Open Office).
- **TXT** : fichiers au format de texte brut, sans mise en forme.
- **DjVu** : format de fichier destiné à l'archivage de documents numériques, créé par AT&T aux États-Unis⁸². Les fichiers sont lus par un logiciel très léger (moins de 3 Mo) et ergonomique. Il permet une compression optimisée en termes de taille et de qualité, et est de plus en plus utilisé par les administrations, mais également par les pirates, notamment pour les bandes dessinées pour lesquelles le visionnage est facilité.
- **HTML** : fichiers inclus dans le corps même d'une page web.
- **IMG** : fichiers dans un format image, généralement en JPEG. Dans le cas des BD, ce format signifie donc un fichier découpé par planches.

	PDF image	PDF texte	Word	TXT	DjVu	HTML	IMG
Total	38,6 %	29,4 %	4,8 %	0,7 %	6,2 %	0,2 %	20,1 %
Livres	38 %	42 %	6,8 %	0,8 %	8,8 %	0,3 %	3,3 %
BD	36,1 %	0 %	0 %	0 %	0 %	0 %	62,9 %

Format des fichiers de livres et BD, hors livres audio.

La très grande majorité des fichiers de livrels sont au **format PDF**, le plus pratique pour une lecture multi-supports. Les formats textes, tous confondus⁸³, représentent **49,9 % des livres piratés**, 35,1 % si on inclut la BD dans le total. Ce sont donc, pour l'essentiel, des fichiers qui ne permettent pas des recherches plein texte, contrairement aux formats numériques que s'approprient à proposer les éditeurs (format PDF texte, ePub...), ce qui est capable de donner un avantage supplémentaire, du moins temporaire, à l'offre légale.

	mp3	wma	m4b
Livres audio	95,1 %	3,3 %	1,6 %

Format des fichiers de livres audio.

En ce qui concerne les livres audio, sans surprise, le **mp3 reste le format standard**.

⁸¹ Dans notre échantillon, on peut repérer ce type d'ouvrages soit, très rarement, par la présence de *watermarks* (tatouages numériques) sur le fichier, soit par le rendu clairement professionnel de ce dernier.

⁸² <http://fr.wikipedia.org/wiki/DjVu>

⁸³ PDF texte + Word + TXT + HTML. DjVu étant un cas à part.

Taille

La taille des fichiers de livrels est beaucoup moins standardisée que celle des fichiers de musique (qui eux sont souvent un format mp3 et compris entre 3 et 5 Mo), de films et de jeux (conçus pour tenir sur un CD de 700 Mo).

	< 1 Mo	1 à 4 Mo	5 à 9 Mo	10 à 49 Mo	50 Mo à 249 Mo	> 250 Mo	Moyenne
Total	14,6 %	17,6 %	6,9 %	34,6 %	20,4 %	5,9 %	77,6 Mo
Livres	20,5 %	27,1 %	10,3 %	28 %	12,9 %	1,2 %	29,7 Mo
BD	4,4 %	0,4 %	2,4 %	55,6 %	30,8 %	6,4 %	79,3 Mo
Livres audio	0 %	0 %	1,6 %	9,3 %	41,2 %	48,3 %	515,7 Mo

Les livres sont en général des fichiers de petite taille comparée à un album de musique ou une vidéo. Toutefois, contrairement à beaucoup d'idées préconçues, l'étude révèle que la taille moyenne est assez élevée : **29,7 Mo** pour les livres, **79,3 Mo** pour les bandes dessinées, et (moins étonnamment) **515,7 Mo** pour les livres audio. Cela traduit, pour les livrels, le grand nombre d'ouvrages illustrés (surtout scolaires et pratiques) scannés au format image qui pèsent beaucoup plus lourd qu'un format texte¹.

La diffusion de livrels par email est donc partiellement un fantasme : elle ne l'est pas dans le cas des livres au format texte, légers et rapidement transférable à plusieurs destinataires en un seul envoi ; en revanche, elle pose problème dans le cas des livres illustrés, des livres audio et des BD, d'autant que ces dernières sont souvent « packagées » c'est-à-dire regroupées en séries complètes.

Les téléchargements, qu'ils soient en *direct download** ou via un outil de peer to peer* ne posent aucun problème de lenteur (temps de téléchargement) à l'heure du haut débit généralisé.

Répartition des releasers

Dans le cadre de l'étude, on appelle « releaser* » l'utilisateur qui crée et diffuse pour la première fois un fichier numérique illégal (et « uploader* » celui qui le diffuse par la suite).

	Rang 1	Rang 2	Rang 3	Rang 4	Rang 5
Releaser	9,5 %	4,3 %	0,8 %	0,3 %	0,2 %

Ce tableau signifie que 9,5 % de l'ensemble des fichiers disponibles proviennent du *releaser** le plus important (de rang 1), 4,3 % du deuxième, etc. Cela permet d'évaluer la part des équipes et des gros utilisateurs dans le piratage des livres² : livres et BD confondus, **deux releasers* concentrent ainsi près de 14 % des créations de fichiers illégaux.**

La qualité moyenne des fichiers de ces deux *releasers** est de 4,7 sur 5, ce qui prouve leur objectif qualitatif et non quantitatif. Dans certains cas, l'impatience d'obtenir un livre non accessible autrement

¹ En considérant seulement les formats textes (PDF texte, Word, TXT et HTML), la moyenne des fichiers est **8 Mo**.

² Dans le domaine de la vidéo, une étude de 2008 a montré qu'environ 1 *release* de film sur 3 était l'œuvre d'un seul pirate ou équipe de pirates : <http://www.slate.com/id/2204367/pagenum/all>

(par exemple un *Harry Potter* pas encore édité en France) prévaut sur la qualité attendue du fichier. Mais, la plupart du temps, le souci de qualité est primordiale et ce d'abord pour des raisons pratiques de lisibilité : il est pénible de lire sur écran ou même d'imprimer un texte de mauvaise qualité.

A noter que la majorité des fichiers n'ont pas de *releaser** clairement identifié.

Analyse des ouvrages

Top 20 des auteurs ayant le plus de titres piratés (livres)

Ce tableau n'indique pas quels sont les auteurs les plus téléchargés illégalement dans l'absolu, mais quels sont ceux qui ont le plus de *titres différents* piratés, d'après nos échantillons.

	Auteur	Titres piratés
1	Gilles Deleuze	13
2	Bernard Werber	11
3	Amélie Nothomb	10
4	Frédéric Beigbeder	7
5	J.K. Rowling	7
6	Michael Connelly	6
7	Sophie Dudemaine	6
8	Jean-Paul Sartre	6
9	Albert Camus	5
10	Daniel Pennac	4
11	Eckhart Tolle	4
12	Harlan Coben	4
13	Michel Foucault	4
14	Isaac Asimov	3
15	Marc Levy	3
16	Ken Wilber	3
17	Paul Ricoeur	3
18	Paulo Coelho ⁸⁶	3
19	Stephenie Meyer	3
20	Ray Bradbury	3

Quelques observations :

❖ **Continuité des best-sellers** : la moitié des auteurs du classement font aujourd'hui partie des auteurs de best-sellers en France (Bernard Werber, Amélie Nothomb, Frédéric Beigbeder, J.K. Rowling, Michael Connelly, Daniel Pennac, Marc Levy, Paulo Coelho, Stephenie Meyer). Leur présence traduit logiquement leur popularité auprès du grand public.

❖ **Surreprésentation de certaines catégories d'auteurs** :

- Les auteurs **philosophes**⁸⁷ : 25 % des auteurs du top 20 sont des philosophes, et c'est l'un d'entre eux qui apparaît en tête de classement. Les explications sont multiples, entre autres : la notoriété de ces auteurs à l'étranger (d'où une volonté des internautes de contourner les coûts d'accessibilité à leurs ouvrages, pas toujours disponibles dans les librairies ou les bibliothèques locales) et, plus généralement, le fort relais universitaire sur le

⁸⁶ Le cas de Paulo Coelho est particulier car il a lui-même diffusé gratuitement les fichiers numériques de ses livres.

⁸⁷ Gilles Deleuze, Jean-Paul Sartre, Albert Camus, Michel Foucault, Paul Ricoeur.

web. Cela souligne également le fait les utilisateurs sont à la recherche de titres indisponibles au format électronique, et qu'il n'y a pas que le grand public (ou l'internaute moyen) qui utilise ses solutions de téléchargement illégal.

- Les auteurs de **science-fiction et fantastique**⁸⁸ : ils représentent également **25 %** du classement. Ils reflètent sans doute les goûts culturels du profil moyen des internautes téléchargeurs⁸⁹.

❖ **Médiatisation d'auteurs ésotéristes** : c'est le cas de Eckhart Tolle et Ken Wilber, deux auteurs de la catégorie « Religion et Esotérisme », qui, dans les réseaux illégaux, est surreprésentée par rapport à sa commercialisation légale (cf. p.48).

Top 20 des titres les plus piratés (livres)

Le classement des titres les plus référencés⁹⁰ recoupe partiellement la liste ci-dessus.

1	Le Sexe pour les nuls	Ruth Westheimer
2	Harry Potter*	J.K. Rowling**
3	Le Grand Livre de cuisine	Alain Ducasse
4	Twilight*	Stephenie Meyer**
5	Les Fourmis*	Bernard Werber**
6	Le Petit Prince	Antoine de St Exupéry
7	L'Alchimiste	Paulo Coelho**
8	Le Seigneur des Anneaux*	J.R.R. Tolkien
9	Les Madeleines salées et sucrées de Sophie	Sophie Dudemaine**
10	Le Larousse des cocktails	Collectif
11	Gâteaux de mamie	Amandine
12	Voyage à motocyclette	Ernesto Che Guevara
13	Nouvelles sous ecstasy	Frédéric Beigbeder**
14	Dieux et Pharaons de l'Egypte ancienne	Claire Lalouette
15	Manuel de survie face aux attentats...	Collectif
16	Fahrenheit 451	Ray Bradbury**
17	Je suis une légende	Richard Matheson
18	Cuisine de l'étudiant	Gudrun Hetzel-Kiefner
19	Millenium*	Stieg Larsson
20	Petits biscuits	Nathalie Helal

**Cumul des différents tomes de la série / ** Fait partie du classement des auteurs ayant le plus de titres piratés*

Quelques constatations :

- **Rééquilibrage en faveur des ouvrages très grand public** et des best-sellers : les auteurs, notamment philosophes, du classement précédent ont beaucoup de titres téléchargés mais chacun en très faibles quantités, ce qui fait qu'ils n'apparaissent pas ici ;

⁸⁸ Bernard Werber, J.K. Rowling, Isaac Asimov, Stephenie Meyer, Ray Bradbury.

⁸⁹ La science-fiction et le fantastique sont parmi les thèmes de prédilection de la culture dite « geek », par ailleurs passionnée d'informatique. On se reportera à l'article de Cory Doctorow : « Science Fiction Is the Only Literature People Care Enough About to Steal on the Internet » (*Locus*, juillet 2006).

⁹⁰ Il s'agit des titres les plus souvent référencés sur les circuits de diffusion illégaux que nous avons explorés, classés par nombre de liens de téléchargements disponibles.

- **Prime à la nouveauté** : certains livres comme *Je suis une légende* de Richard Matheson bénéficie d'un effet de mode lié à la sortie au cinéma du film du même titre (décembre 2007, sortie DVD en France en juin 2008). De même pour Pierre Desproges dont la célébration des 60 ans de sa naissance ont donné lieu à de nombreuses rééditions au printemps 2009.
- **Des volumes de téléchargement encore faibles et partiellement comptables** : les chiffres ne figurent pas dans le tableau ci-dessus car ils ne peuvent être considérés qu'à titre indicatif : ils ne représentent qu'une vision partielle des téléchargements de chaque fichier, le calcul exhaustif étant par définition irréaliste (notamment parce que certains agrégateurs ne fournissent pas de statistiques et que celles-ci sont quasiment impossibles sur le réseau ed2k.)

Titre	Téléchargements
<i>Le Sexe pour les nuls</i>	58 820
<i>Harry Potter</i>	39 453
<i>Le Grand livre de cuisine</i>	34 022
<i>Twilight</i>	19 159
<i>Les Fourmis</i>	5 202

Téléchargements d'après les principaux agrégateurs de torrent et de direct download

Il faudrait appliquer aux résultats trouvés un coefficient multiplicateur difficile à déterminer : notre hypothèse est que ce coefficient est un minimum de 2 ou 3, ce qui porterait les titres les plus téléchargés autour de 150 à 200 000 téléchargements. Quelle que soit la méthode de calcul, ces chiffres sont négligeables au regard des téléchargements de musique ou de films¹.

Top 20 des éditeurs ayant le plus de titres piratés (livres)²

Ce tableau n'indique pas quels sont les éditeurs les plus piratés dans l'absolu, mais quels sont ceux qui ont le plus de titres différents piratés.

1	Gallimard*	11	Seuil
2	Dunod	12	Larousse
3	Hachette*	13	Descartes Editions
4	Eyrolles	14	Minerva
5	Editions de Minuit	15	Marabout
6	PUF	16	First
7	Albin Michel	17	Multisim Editions
8	Flammarion	18	Assimil
9	EDP Sciences	19	Asmodée Editions
10	Ellipses	20	La Découverte

* dont Jeunesse

Le classement reflète une double réalité :

- Les gros éditeurs sont aussi les plus piratés ;
- Certaines catégories d'édition sont très présentes : c'est le cas des **STM*** et des **jeux de rôles**¹. Les éditeurs de STM* disposent, il faut le noter, de catalogues de livres numériques

¹ Films les plus téléchargés en 2009 en torrent : *Watchmen* (16 906 452 téléchargements), *The Curious Case of Benjamin Button* (13 133 137) et *Yes Man* (13 038 364). Source : BigChampagne.

² Le classement ne tient pas compte des rééditions au format poche. Les éditeurs les plus piratés dans ce domaine sont, dans l'ordre : Le Livre de Poche, Pocket, J'ai Lu et 10/18.

particulièrement bien fournis. Une bonne partie des fichiers piratés semblent l'être, d'après notre analyse, depuis des plateformes légales de téléchargement.

Top 12 des éditeurs ayant le plus de titres piratés (BD)

1	Delcourt
2	Dargaud
3	Dupuis
4	Glénat
5	Soleil Productions
6	Vents d'Ouest
7	Fluide Glacial
8	Casterman
9	Albin Michel
10	Cap
11	International Press Magazine
12	Les Humanoïdes Associés

Le classement contient deux éditeurs marginaux (Cap et International Press Magazine), dont les publications reflètent la surreprésentation des bandes dessinées pour adultes (pornographiques). Il est à comparer avec les parts de marché des éditeurs :

1	Soleil Productions	4,8 %
2	Delcourt	4,6 %
3	Dupuis	4 %
4	Casterman	3 %
5	Glénat	2,9 %
6	Dargaud	2,4 %
7	Le Lombard	1,9 %
8	Les Humanoïdes Associés	1,5 %
9	Bamboo	1,3 %
10	Vents d'Ouest	1,2 %
11	Futuropolis	1,1 %
12	Fluide Glacial	1 %

Parts de marché en nombre de titres édités en 2008, hors éditeurs de mangas. Source : Livres-Hebdo/ACBD².

A cette liste il convient d'ajouter une analyse spécifique du marché du **manga** (scantrad*), qui fut l'un des premiers à subir le piratage dès le milieu des années 1990. L'ampleur du phénomène est d'autant plus importante qu'il est corrélé au succès du genre : en 2008, le manga représentait 40,5 % des nouveautés en bandes dessinées en France³. Il est toutefois difficile d'obtenir des chiffres réalistes car beaucoup d'ouvrages sont consultés directement sur le web et non téléchargés. A titre d'exemple, le classement d'un des principaux sites concernés est le suivant :

- 1) *Bleach*, avec 335 000 visites
- 2) *Naruto*, 294 000 visites

¹ Ce qui fait entrer dans le classement des petits éditeurs inconnus du grand public comme Descartes Editions, Multisim Editions et Asmodée Editions.

² http://www.acbd.fr/images/stories/acbd_bilan_2008.pdf

³ Source : ACBD © Gilles Ratier, secrétaire général de l'ACBD.

3) *One Piece*, 268 000 visites.

Ces chiffres s'entendent tous albums confondus, les trois titres cités étant des séries.

Année de publication

Un même ouvrage, piraté dans une réédition des années 1990, peut avoir été originellement publié chez un autre éditeur dix ans plus tôt. Les chiffres qui suivent ont donc une valeur indicative. Ils tiennent compte, dans la mesure du possible, de la première date de publication et non pas des rééditions ultérieures.

	1900-1950	1951-1970	1971-1990	1991-2000	2001-2005	2006-2007	2008-2009
Livres	1,1 %	4 %	10,5 %	21,9 %	36,4 %	18,2 %	7,9 %
BD	0 %	0,2 %	4,1 %	32,8 %	29,8 %	17,4 %	15,7 %

Plus de **2 ouvrages piratés sur 3, livres comme bandes dessinées, datent de moins de 10 ans**. 26,1 % des livres et 33,1 % des BD ont été publiés il y a moins de 4 ans. Le piratage ne touche donc pas seulement un fond de catalogue épuisé, il concerne les productions récentes.

Toutefois on constate que ces productions récentes concerne pour plus de la moitié une période allant de 1991 à 2005 : le partage se fait moins sur les nouveautés (contrairement aux films et à la musique) que sur des titres assez établis, très diffusés. En matière de livres pirates, les internautes semblent chercher davantage l'ouvrage de valeur sûre ou répondant à un besoin (ou une curiosité) immédiat, que la nouveauté à proprement parler.

Origine des ouvrages

Rappelons que nous n'étudions que les œuvres éditées par des éditeurs français, en français (traduites ou non), et que nous excluons les traductions d'œuvres françaises à l'étranger.

	France	Etranger (traductions)
Livres	78,1 %	21,9 %
BD	90,4 %	9,6 %

Le tableau est à comparer avec la part des traductions dans la production éditoriale : 14 % selon les chiffres de *Livres-Hebdo*/Electre en 2008. Par ailleurs 99 % des ouvrages traduits et piratés proviennent des versions officielles parues chez les éditeurs autorisés. Seules quelques rares exceptions (certaines versions des romans de J.K. Rowling par exemple) sont le fruit de traductions officieuses réalisées par des fans, généralement impatients de diffuser les textes avant leur sortie en France¹.

¹ A cet égard on peut considérer que la démultiplication, ces dernières années, des campagnes marketing pour quelques titres best-sellers, à grands renforts de *teasers* et précommandes, entraîne chez les lecteurs des pulsions consuméristes qui les poussent à assouvir leur besoin grâce au piratage.

Genres (livres)

Comprendre quels sont les types d'ouvrages piratés permet de mieux appréhender certains besoins des internautes en matière d'offre numérique, même s'il n'est pas évident que le téléchargement illégal se superpose avec exactitude à l'achat au format numérique. En guise de comparaison, le tableau indique la répartition des ventes de livres en France en 2008.

	Fichiers illégaux	Ventes papier légales*
Essais, documents	28,8 %	8,7 %
Romans	27,1 %	25,6 %
Pratique	26,7 %	10,8 %
Scolaire (et parascolaire)	12,9 %	12,1 %
Jeunesse	2,2 %	17,4 %
Beaux-livres	0,9 %	2 %
Poésie, théâtre	1,4 %	0,6 %

*Part des exemplaires vendus (source : SNE, 2009)¹

On constate principalement :

- ❖ Une **surreprésentation des essais et documents** : celle-ci est à relier à la forte présence de la philosophie (cf. p.42 et 48)
- ❖ Une **surreprésentation des ouvrages pratiques**, au sein desquels les livres de cuisine constituent environ 34 % des fichiers. (Ce qui témoigne pour le moins de l'importance de la gastronomie dans les hobbies des internautes !)

Ces deux catégories d'ouvrages partagent une même caractéristique : ils sont facilement consultables sur des supports numériques, car ils ne demandent pas une lecture linéaire et répondent à un besoin immédiat (scolaire, professionnel ou personnel) en termes d'information.

La faible présence du théâtre et de la poésie est proportionnelle à sa marginalité en termes d'exemplaires vendus (et de chiffre d'affaires) dans le livre papier. En ce qui concerne la poésie, par ailleurs, il faut préciser que l'on peut trouver des milliers de poèmes sous droits, souvent isolés, publiés illégalement sur des forums et des pages personnelles, directement consultables sans téléchargement.

Enfin la sous-représentation du beau-livre s'explique par la difficulté de scanner les formats originaux. Celle de la littérature jeunesse, par l'inadéquation du support électronique par rapport aux usages des plus jeunes, et la moyenne d'âge des internautes téléchargeurs.

¹ Les catégories de fichiers recourent la segmentation éditoriale du SNE de façon simplifiée. Ainsi la catégorie « bande dessinée » (9,4 % des exemplaires vendus en 2008) n'est pas considérée ici, car elle est traitée à part.

Thèmes (livres)

Les thèmes figurant ci-dessous correspondent aux principales catégories observées. La rubrique « Autres » regroupe tout le reste.

STM*	18,1 %
Cuisine	13,6 %
Philosophie	12,8 %
Jeux	11,4 %
Science-fiction et fantastique	11,4 %
Informatique	6,4 %
Religion et ésotérisme	6,1 %
Policier	5,3 %
Psychologie	3,2 %
Biographie	2,5 %
Adulte	1,1 %
Santé	0,7 %
Tourisme	0,4 %
Humour	0,3 %
Autres	6,7 %

On remarque notamment une forte représentation des ouvrages de **STM***, de **cuisine**, de **philosophie**, de **jeux** (essentiellement les jeux de rôles) et la **littérature de science-fiction et fantastique**. Ces 4 catégories représentent à elles seules **67,3 %** des livres piratés de notre échantillon.

Seules quelques catégories sont comparables avec les chiffres de l'édition donnés par le SNE (qui n'opère qu'un regroupement thématique partiel). C'est le cas de la catégorie « Religion et ésotérisme » (1,5 % des exemplaires vendus) et la catégorie « Policiers » (4,7 %)¹.

L'informatique (6,4 %) semble sous-représentée au regard de sa quasi omniprésence dans les résultats de recherche de livrels illégaux. Cela s'explique par deux facteurs : d'une part, la méthode d'échantillonnage (p.24) et d'autre part, le fait qu'une grande partie des fichiers d'informatique ne soient pas des livres à proprement parler mais des ouvrages dérivés : manuels d'utilisation, documents créés par des utilisateurs, etc.

Prix moyen

Le prix de l'offre légale de livres, papier comme numérique, est cruciale pour comprendre le piratage. Le tableau ci-dessous indique le prix en offre légale papier des ouvrages piratés (il tient seulement compte des ouvrages encore disponibles neufs à la vente²).

	< 5 €	5 à 10 €	11 à 20 €	21 à 30 €	> 30 €	Moyenne
Livres	11,8 %	36,7 %	25,9 %	15,1 %	10,5 %	15,7 €
BD	2,6 %	34,2 %	57,9 %	3,4 %	1,9 %	12,4 €

¹ Source : SNE, 2008.

² Dans de nombreux cas, les ouvrages sont disponibles à la vente parallèlement en édition grand format et poche. Nous n'avons dans ce cas retenu que le prix au format poche, considérant que c'était à ce prix, le plus bas possible, qu'il fallait comparer l'offre illégale.

Le prix moyen du livre papier en vente légale des ouvrages piratés, toutes catégories confondues, s'élève à **15,7 € pour les livres** et **12,4 € pour les bandes dessinées**.

Prix moyen par catégorie :

Beaux-livres	28,3 €
Pratique	23,8 €
Scolaire	23 €
Essai	16,7 €
Jeunesse	15,9 €
Roman	11,4 €
Théâtre	9,1 €
Nouvelles	8,6 €
Poésie	5,6 €

Le prix du livre papier, s'il est peu influent concernant une grande partie des ouvrages de fiction achetables au format poche pour moins de 10 euros, est-il un facteur important de la décision du piratage d'ouvrages pratiques et scolaires, dont la moyenne cumulée s'élève à 23,5 € euros ? La question reste ouverte.

Disponibilité des ouvrages

Il est enfin intéressant de faire ressortir le **nombre d'ouvrages piratés non disponibles à la vente** :

Livres	25,6 %
BD	31,4 %

Ces taux reflètent l'une des motivations des téléchargeurs : avoir accès à des ouvrages épuisés ou non disponibles. Ils soulignent **l'importance pour les éditeurs de développer une offre légale la plus exhaustive possible**. Cependant, puisque la très grande majorité des titres sont disponibles à la vente, d'autres critères de motivations sont ici à l'œuvre : le prix, soit, mais surtout l'accès à un contenu sous sa forme numérique.

Classements

Après examen de la convergence des tops 50 (2008) de *Livres-Hebdo* et de l'offre disponible au téléchargement illégal de nos modèles, voici ce qu'il apparaît :

	Romans	Essais	Poches	Jeunesse	Pratique	BD	Total
Titres en téléchargement illégal	4 (8 %)	1 (2 %)	6 (12 %)	4 (8 %)	3 (6 %)	11 (22 %)	27 (9 %) Hors BD:7,2%
Exemplaires vendus (1^{er} et dernier du top 50)	512 800 43 900	231 200 29 100	475 400 78 900	177 200 28 200	140 900 27 600	495 400 45 200	

Source : Top 50 des ventes 2008, *Livres-Hebdo*

La catégorie « Romans » est à 7 (14 %) si l'on y considère la série *Twilight* de Stephenie Meyer (classée dans la catégorie « Jeunesse » par *Livres-Hebdo*). La catégorie « Poches » est la plus représentée, ce qui traduit la forte présence en téléchargement illégal des succès de librairie récents et reconnus, plus que des toutes dernières nouveautés.

Les titres concernés sont les suivants :

Romans	<i>Millenium</i> , tome 1 (Stieg Larsson) <i>Millenium</i> , tome 2 (Stieg Larsson) <i>Millenium</i> , tome 3 (Stieg Larsson) <i>L'Élégance du hérisson</i> (Muriel Barbery)
Essais	<i>L'Instinct de mort</i> (Jacques Merisne)
Poches	<i>Les Enfants de la liberté</i> (Marc Levy) <i>Et après</i> (Guillaume Musso) <i>Je voudrais que quelqu'un m'attende quelque part</i> (Anna Gavalda) <i>Les Bienveillantes</i> (Jonathan Littell) <i>Et si c'était vrai</i> (Marc Levy) <i>Le Petit Prince</i> (Antoine de Saint-Exupéry)
Jeunesse	<i>Twilight</i> , tome 1 (Stephenie Meyer) <i>Harry Potter et les reliques de la mort</i> (J.K. Rowling) <i>Twilight</i> , tome 2 (Stephenie Meyer) <i>Twilight</i> , tome 3 (Stephenie Meyer)
Pratique	<i>La Méthode simple pour en finir avec la cigarette</i> (Allen Carr) <i>Pain maison</i> (Collectif/Marabout) <i>Divines verrines</i> (Maya Barakat-Nuq)

Classement par ordre décroissant de ventes papier.

Ces données ne recouvrent que les livres échantillonnés par l'étude, ils sont toutefois assez représentatifs de la globalité du piratage des best-sellers : selon nos observations des fichiers piratés en dehors de notre échantillon, à l'exception de la catégorie BD, on n'y trouve que peu des autres titres du top 50, moins de 10 par catégorie. En tenant compte de ce différentiel, nous estimons que le **taux de piratage global des livres (hors BD) du top 50 des ventes** passe de 7,2 % à environ **8 %**.

Nous avons choisi de comparer l'offre illégale avec les catalogues des principaux vendeurs de livrels en France¹ :

- ❖ Mobipocket : environ 100 000 titres ;
- ❖ Numilog (racheté par Hachette en 2008) : distributeur des principales maisons d'éditions françaises et partenaire de Fnac.com (30 000 références selon *Livres-Hebdo*, 28/09/2009).
- ❖ Immatériel : environ 2 400 titres ;
- ❖ I-Kiosque : environ 2 000 titres ;
- ❖ Epagine : environ 1 200 titres ;
- ❖ Pour les livres audio, l'agrégateur Audible (groupe Bertelsmann) et l'éditeur-libraire Livraphone ;
- ❖ En ajoutant quelques revendeurs plus modestes, comme Digibidi pour la bande dessinée (environ 120 titres), LeKiosque.fr (environ 120 titres).

Il en résulte que **94,9 % des livres piratés ne disposent pas d'une offre numérique légale**. Les livres audio et, dans une moindre mesure, les romans, sont les mieux servis à cet égard :

Livres	5,1 %
Romans	8,1 %
Essais	3,8 %
Pratique	2,9 %
Scolaire	6,9 %
BD	1,2 %
Livres audio	18,4 %

Sur cette offre numérique légale, le prix moyen des ouvrages est le suivant² :

Livres	26 €
Livres audio	16,9 €

On observe donc un décalage important, concernant les livres, entre le prix proposé par les éditeurs (26 €) et le prix moyen que les lecteurs se déclarent prêts à payer, qui s'élève à **6 €** selon une étude de GFK en février 2009³. On peut voir dans ce décalage l'une des causes essentielles du piratage. Toutefois, il faut relativiser l'argument du prix dans la mesure où, d'une part, le niveau de piratage est encore faible et, d'autre part, les efforts des pirates sont encore concentrés sur les best-sellers accessibles à prix réduit au format poche.

¹ Offre disponible au moment de l'étude, à l'été 2009. Etant donné l'évolution rapide en la matière, et le lancement d'Eden Livres (principal concurrent de Numilog) en septembre 2009, des mises à jour seront donc nécessaires. A noter que les différents catalogues ne sont pas toujours exclusifs et qu'il y a donc de forts recoupements entre eux.

² Les BD ne sont pas incluses dans le tableau car leurs versions numériques légales sont encore trop peu nombreuses pour en tirer un résultat pertinent.

³ www.sne.fr/pdf/Tendances_activite_editoriale2008-2009.pdf

Synthèse des résultats

- ❖ Le téléchargement illégal des livres sur Internet n'est pas nouveau, mais il semble décoller depuis le **milieu de l'année 2008** à travers le monde. Il reste toutefois un phénomène particulièrement mineur au regard des autres formes de piratage en ligne.
- ❖ L'évaluation du nombre de livres disponibles en téléchargement illégal reste complexe. Mais on peut l'estimer, à l'été 2009, de **4 000 à 6 000 titres différents**, dont 3 000 à 4 500 bandes dessinées. Soit **moins de 1 % des titres disponibles légalement** au format papier.
- ❖ Pirater un livre en le scannant depuis sa version papier représente un travail fastidieux de **6 à 10 heures** minimum pour un résultat de bonne qualité.
- ❖ Pourtant, **77 % des ouvrages atteignent ce niveau de qualité**. 1 ouvrage piraté sur 7 l'est par l'une des deux principales équipes de pirates, qui privilégient la qualité à la quantité.
- ❖ La **taille moyenne d'un livre piraté est de 29,7 Mo**. Ce poids assez élevé traduit l'importance du nombre d'ouvrages illustrés, notamment scolaires et pratiques, en téléchargement illégal.
- ❖ Le réseau peer to peer **eDonkey** ainsi que le **direct download** sont les circuits de diffusion les plus utilisés pour les livres. Mais, contrairement à la musique, à la vidéo et aux logiciels, la faible quantité de fichiers disponibles et d'utilisateurs pour les partager rend leur accès **relativement difficile**.
- ❖ **Gilles Deleuze, Bernard Werber et Amélie Nothomb** sont les auteurs ayant le plus de titres disponibles en téléchargement illégal.
- ❖ Les livres les plus souvent indexés et partagés dans les différents réseaux illégaux sont **Le Sexe pour les nuls**, la série **Harry Potter** de J.K. Rowling, le **Grand Livre de cuisine** d'Alain Ducasse et la série **Twilight** de Stephenie Meyer.
- ❖ Les trois éditeurs ayant le plus de titres piratés sont **Gallimard, Dunod et Hachette**.
- ❖ Le manque de disponibilité des ouvrages en offre légale ne semble être qu'une cause partielle du piratage : **3 livres piratés sur 4 sont disponible à la vente papier**. Mais **94,9 % des livres piratés ne disposent pas d'une offre numérique légale**. Le piratage concerne les **parutions contemporaines** : 2 ouvrages piratés sur 3 ont été publiés il y a moins de 10 ans, mais **pas les nouveautés en rayon** : seul 1 ouvrage piraté sur 4 a été publié il y a moins de 4 ans.
- ❖ Les catégories « **essais** », « **romans** » et « **livres pratiques** » représentent chacune un peu plus de 25 % des livres piratés.
- ❖ Seuls **8 %** des livres du **top 50 des ventes papier en librairie** sont disponibles en téléchargement illégal.
- ❖ Concernant les thèmes, on constate une très forte présence des **STM***, de la **philosophie**, de la **cuisine**, de la littérature de **science-fiction** et **fantastique** et de **l'informatique**.

Conclusion

La problématique du numérique pour les livres relève, en partie, d'un choc culturel entre deux univers extrêmement différents, les « literati » vs les « technorati »¹. Comme le rappelle Liisa McCloy-Kelley, VP Director Content Management de Random House : « Nous avons tendance à considérer la lecture comme une activité à part, mais elle fait aussi partie de la convergence des médias numériques². » Le téléchargement illégal des livres, à la suite du téléchargement des autres contenus culturels numériques, est une réalité aujourd'hui, même si encore très limitée, en volume comme en nombre d'utilisateurs.

Faut-il avoir peur du piratage ?

Les personnes qui ne croient pas dans l'avenir du livre numérique ne sont-elles pas souvent celles qui avancent l'argument du piratage massif ? Cette apparente contradiction tiendrait à la nature de la courbe d'expérience du marché du livre, qui n'a pas encore atteint son « tipping point »³.

Dans une première période, que nous vivons actuellement, les conséquences positives ou négatives du téléchargement illégal des livres sont faibles, car les volumes échangés le sont également.

Dans une deuxième période, lorsque les nouveaux usages de lecture numérique auront su conquérir le grand public (notamment grâce aux nouvelles générations de liseuses), il est probable que le marché du livre soit confronté aux mêmes problématiques que le marché de la vidéo et de la musique. Selon cette analogie et pour reprendre les 8 points présentés par Kevin Kelly (cf. p.16), l'offre légale devra alors peut-être répondre à un certain nombre de critères susceptibles de la rendre plus intéressante que l'offre piratée. Il est évident qu'une initiative comme celle d'Amazon, largement controversée, de supprimer directement les livres litigieux sur les Kindle de ses utilisateurs a apporté de l'eau au moulin des pirates⁴. « Si j'ai plus de problèmes avec l'offre légale qu'en piratant, alors autant pirater » : voilà brièvement résumé le raisonnement qui a contribué à discréditer l'usage des DRM* dans la musique, et qui, pour certains, n'incite pas à les maintenir pour le livre. Les DRM* ont par ailleurs montré leurs limites en matière d'interopérabilité.

Quelques recommandations aux éditeurs

Sans exagérer l'importance du piratage, il ne faut pas non plus l'ignorer. On pourra donc conclure cette étude sous forme de recommandations aux éditeurs. Tout d'abord en les incitant à développer leur

¹ « Il y a un clash entre ceux qu'on pourrait appeler les technorati et les literati », écrit Bob Thompson dans le Washington Post (...) Les technorati sont enthousiasmés par la façon dont l'Internet et les ordinateurs révolutionnent le monde des livres. Les literati craignent que, au milieu de cette ferveur révolutionnaire, des institutions et des valeurs essentielles ne soient guillotonnées. », Jeff Gomez : *Print is dead*, Macmillan, 2008, p.33 (nous traduisons).

² « We tend to think about reading as separate from other media, but it is becoming part of a larger digital media convergence. », conférence « Digital Book », New York, 12/05/2009.

³ Point de basculement rapide à partir duquel un produit ou un usage est adopté par le grand public.

⁴ Rappelons qu'en juillet 2009, Amazon a, pour des raisons juridiques liées à la protection du droit d'auteur (les ouvrages ayant été diffusés sur le site d'Amazon par une société qui n'en détenait pas les droits), supprimé deux ouvrages de George Orwell directement sur le Kindle de ses abonnés, sans avertissement.

catalogue numérique, condition *sine qua non* pour concurrencer l'offre illégale, et à proposer des livrels de qualité à des prix attractifs. A ces soucis de qualité de l'offre doit-on associer un changement de logique économique et du métier d'éditeur: l'éditeur du futur serait-il amené à davantage développer des services liés aux textes et à la lecture ? A cet égard, certains ont déjà tenté de tirer parti de la création de communautés d'e-lecteurs et de l'utilisation des différents outils de réseaux sociaux.

Pour résumer, nous suggérons aux éditeurs de :

- Développer leur catalogue numérique, l'asseoir sur des modèles économiques innovants et des services spécifiques.
- Développer une stratégie de référencement dans le but de faire apparaître en premier l'offre légale sur les moteurs de recherche, grâce à des outils et des compétences techniques appropriés.
- Développer les alertes et la surveillance autour de leurs titres, afin de réguler, si besoin, les téléchargements, mais surtout pour mieux appréhender la demande et l'attractivité de leurs livres auprès du public.
- Mettre en œuvre, avec les collectivités territoriales, le plus rapidement possible, des formations professionnelles destinées à l'édition indépendante afin d'éviter la création d'une « fracture » numérique entre les « grands » éditeurs et la petite édition.
- Ne pas mener une guerre contre les lecteurs. Faut-il considérer que le téléchargement illégal ou gratuit d'un livre signifie automatiquement une vente perdue ? Cela n'est en tout cas pas ce que mettent en avant les expériences « libres » comme celles, parmi d'autres, des éditions de l'Eclat ou encore du label Zone de La Découverte.
- Plus largement, réfléchir, en concertation avec les autres maillons de la « chaîne » du livre et surtout avec les auteurs, à la meilleure manière de faire respecter leurs droits et leur rétribution ainsi qu'à la valeur d'usage des livres. Le danger pour les éditeurs n'est peut-être pas tant le téléchargement illégal que le remodelage économique du secteur, au profit de ceux qui, comme l'a avancé le rapport Patino, ne détiennent pas les droits pour l'instant (par exemple, les opérateurs télécoms).

Le piratage des livres sur Internet doit être abordé avec calme et pragmatisme. Il n'est ni un problème négligeable, ni un cataclysme économique. C'est une réalité dont on doit cerner les causes et les conséquences. Par exemple, la contraction des budgets des ménages dans le domaine de la culture (notamment dans le domaine du livre par rapport à d'autres divertissements) aura-t-elle pour conséquence de favoriser le développement d'offres gratuites, légales ou non ?

On ne peut prédire quelle sera l'ampleur du téléchargement illégal dans les années à venir, mais il est certain que, à défaut d'offre légale attractive, adaptée aux usages des internautes, c'est vers des sources illégales que se tourneront ces derniers.

Cette absence de prévisibilité plaide d'ailleurs pour la reconduction périodique du type d'investigation que nous avons menée et dont les principaux résultats ont été ici détaillés ; et ce, de manière à étudier la nature de l'offre numérique illégale de livrels et son éventuelle indexation aux évolutions commerciales et technologiques.

La méthode adoptée ici pourrait être complétée d'une recherche ciblée sur un corpus de titres précis, d'après les classements de meilleures ventes ou d'après certains catalogues d'éditeurs qui souhaiteraient participer à l'expérimentation. Et c'est surtout une étude sur les usages, les comportements et les demandes des téléchargeurs, mais aussi des pirates, qui serait de nature à éclairer les mécanismes du piratage et ces questions cruciales : les gros téléchargeurs sont-ils de gros lecteurs ? le piratage concerne-t-il plus les best-sellers que la petite édition ? quelles sont les conséquences réelles sur les ventes ?

Lexique

binaire	fichier qui ne peut être lu au format texte (par exemple : un format image). Les fichiers échangés dans les newsgroups le sont sous cette forme, découpés en plusieurs parties qui seront réassemblées par la suite par les téléchargeurs.
direct download (DDL)	fichier disponible en téléchargement direct depuis un lien vers un site de stockage spécialisé (les plus connus étant MegaUpload et RapidShare).
darknet	réseau privé de partage de fichiers, dont l'accès est restreint à certains utilisateurs.
DRM	acronyme de « Digital Rights Management », qui recouvre les procédés de protection technique des droits pour les contenus numériques. Les éditeurs de livres français n'ont pas signé les accords Olivennes et se réservent le droit d'utiliser des DRM.
ebook	traduit par « livre électronique » ou « livrel », il s'agit de la version numérique d'un livre, consultable depuis un ordinateur, un appareil portable ou un lecteur spécialisé appelé <i>reader</i> .
eDonkey (ed2k)	système de partage de fichiers en <i>peer to peer</i> , dont le logiciel le plus populaire est eMule.
fake	fichier dont le nom ne correspond pas à son contenu. Les <i>fakes</i> sont créés soit par des personnes souhaitant profiter de la notoriété de certains contenus pour diffuser le leur, soit par les producteurs eux-mêmes afin de polluer le téléchargement illégal.
leecher	(en anglais : « sangsue ») utilisateur de <i>peer to peer</i> qui télécharge sans partager ses fichiers.
peer to peer (P2P)	abrégié en P2P, il s'agit d'un système décentralisé d'échanges de contenus numériques, d'ordinateur à ordinateur (par opposition au modèle client-serveur).
reader	traduit par « liseuse », il s'agit d'un appareil de lecture d'ebooks. Les leaders du marché sont aujourd'hui Sony (Sony Reader) et Amazon (Kindle). On appelle également ainsi les applications pour ordinateur ou appareils portables comme l'iPhone, offrant une fonction similaire (comme Stanza ou eReader).
release	diffusion sur les réseaux d'un fichier numérique, en général illégal. On parle de « release group » à propos de l'équipe qui effectue cette diffusion.
scantrad (scanlation)	manga numérisé, traduit (généralement depuis l'anglais ou le japonais), et diffusé illégalement et la plupart du temps gratuitement sur Internet. Il s'agit en général de mangas non disponibles à la vente en France.
seeder / seed	(en anglais : « semeur ») utilisateur de <i>peer to peer</i> qui partage du contenu (des copies complètes de fichiers), par opposition au <i>leecher</i> . Un « seed » est un utilisateur partageant un fichier complet.
STM	ouvrages de sciences, technique et médecine.
team	équipe de création et de diffusion de fichiers pirates, parfois structurée de façon quasi professionnelle.
téléchargeur	personne qui télécharge des fichiers sur Internet.
torrent	désigne soit le protocole d'échanges en <i>peer to peer</i> BitTorrent, soit le fichier de métadonnées contenant les informations sur le fichier à télécharger, soit encore le fichier lui-même.
tracker	désigne à la fois le fichier de référence d'un contenu diffusé en torrent ainsi que les sites indexant ces fichiers.
warez	tout contenu protégé et diffusé illégalement sur les réseaux.

Annexes

Annexe 1 : Top 50 des livres les plus téléchargés sur Internet en Allemagne en 2007

Rang	Nom du fichier
1	Der.Brockhaus.Multimedial.Premium.20
2	Glasklar 3D Naturwissenschaften v7.0
3	Microsoft Encarta 2007 Lernen und Wissen
4	Perry Rhodan Silberedition 12 - Der Anti
5	Langenscheidt Vokabeltrainer Englisch
6	Glasklar Lexikon 2007 Informatik
7	Digitale Bibliothek - Quellen Philosophie
8	Perry Rhodan Silber
9	digitale.bibliothek.bd.02.philosophie
10	NFPA - Collection fire building mechanik
11	(ebook - engineering - pdf) - Fundamental
12	Rmc - Pmp Exam Prep
13	Bauch
14	Tomtom 5.2 Smartphone Deutschland
15	Pdf Socrate Platon Aristote Kant Heg
16	Duden.Die.Deutsche.Rechtschreibung.2006
17	Duden.Korrektor.PLUS.3.5.1
18	Assimil - English for Business World
19	Fansadox Collection 73 - The Convent
20	Friedrich Nietzsche
21	Duden Die Deutsche Rechtschreibung 2
22	Damnation City
23	Medizin Pschyrembel Klinisches Wb 20
24	Langenscheidt Basic German Grammer
25	O'reilly - Learning Uml
26	ELT Oxford University Press - English
27	Lexikon Sicherheit und Gesundheit
28	Anthony Robbins - the time of your life workbook
29	McCarthy- O'Dell_ Advanced- vocabula
30	TomTom Western and Central Europe v6
31	Delta Next Generation Toefl Test
32	(Electrical Engineering) Prentice Hall
33	Prentice Hall - Digital Modulation & Coding
34	Langenscheidt Woerterbuch Technik
35	Paulo Coelho - Sei Wie Ein Fluss
36	WOD - Vampire - The Masquerade
37	IGI.Global.Encyclopedia.of.Mobile.Computering
38	2_Duden_Korrektor_PLUS_v4.0
39	Prentice Hall - W.Stallings - Data and Computer Communications
40	Note On Oxford Practice Grammar International
41	Langenscheidt - E-Fachwoerterbuch Technik
42	Digitale Bibliothek S014 Lexikon
43	Teach Yourself Visual Studio Net In 21 Days
44	Perry Rhodan - 13 - Atlan - Die Rebe
45	The Oxford-Duden Pictorial English Dictionary

46	libreremo - H.Aufderstrabe – W.Bonzl
47	Brian Tracy - The Psychology of Achievement
48	Der Grosse National Geographic Photocolelction
49	Communication Systems 4th Edition
50	Oxford Picture Dictionary

Source : Ipoque Internet Study 2007

Annexe 2 : 50 résultats les plus récents d'une recherche de livrels sur Mininova (torrent)

Date d'ajout	Description du fichier
10-juil-09	Supply Chains - A Manager's Guide in Ebooks
10-juil-09	Interactive Curve Modeling release in Ebooks
10-juil-09	QuickBooks 2009 All-in-One For Dummies in Ebooks
10-juil-09	Piano For Dummies in Ebooks
10-juil-09	Page -A- Minute Memory Book in Ebooks
09-juil-09	Beginning Microsoft Word Business Documents - From Novice To Professional (2006) - tk in Ebooks
09-juil-09	Manipulieren Aber Richtig Die acht Gesetze der Menschenbeeinflussung in Ebooks
09-juil-09	The Execution of Willie Francis (2008) - in Ebooks
10-juil-09	Lynn Viehl - Darkyn Series in Ebooks
10-juil-09	Richelle Mead - Vampire Academy Series Books 1-3 in Ebooks
09-juil-09	Patrick O'Brian - The Aubrey/Maturin Series in Ebooks
09-juil-09	Rick Sternbach - Star Trek: The Next Generation Technical Manual in Ebooks
10-juil-09	Twilight Saga, The Host, Midnight Sun Excerpt - Stephanie Meyer in Ebooks
10-juil-09	Beddor - The Looking Glass Wars & Seeing Redd in Ebooks
10-juil-09	Alice's Adventures in Wonderland & Through the Looking-Glass in Ebooks
10-juil-09	Dan Brown - Angels & Demons in Ebooks
10-juil-09	Inheritance Cycle - Eragon, Eldest, Brisingr by Christopher Paolini in Ebooks
10-juil-09	House of Night Series Books 1-5[.] in Ebooks
10-juil-09	Charlaine Harris-Sookie Stackhouse 1-7[.] in Ebooks
09-juil-09	Religious Motivation & Origins Buddhism - in Ebooks
10-juil-09	Extreme Exploits - Advanced Defenses Against Hardcore Hacks (2005) - in Ebooks
10-juil-09	Cryptography And Network Security, 4th Edition (2005) - in Ebooks
10-juil-09	1 commentsDigital Photography Hacks - 100 Industrial-Strength Tips & Tools (2004) - in Ebooks
10-juil-09	Excel Programming Weekend Crash Course (2003) - in Ebooks
10-juil-09	Google Search And Tools In A Snap (2006) - in Ebooks
10-juil-09	How Buddhism Began 2nd Ed - in Ebooks
10-juil-09	Access Forms & Reports For Dummies (2006) - in Ebooks
10-juil-09	How To Make Love All Night Drive A Woman Wild - in Ebooks
10-juil-09	100 Weird Facts About the Human Body pdf in Ebooks
10-juil-09	Laptops Just the Steps For Dummies release in Ebooks
10-juil-09	Picturing Machines 1400-1700 release in Ebooks
10-juil-09	DNS On Microsoft Server 2003 in Ebooks
09-juil-09	James Grippando - Got the Look in Ebooks
10-juil-09	Google Images Trick in Ebooks
09-juil-09	The PHP Anthology, Volume I Foundations-Volume II Applications - in Ebooks
09-juil-09	CSS - The Missing Manual (2006) - in Ebooks

10-juil-09	Wiley Wireless Broadband Conflict and Convergence Nov 2008 eBook-BBL in Ebooks
10-juil-09	The ULTIMATE Tesla Coil Design and Construction Guide, 1st Edition in Ebooks
10-juil-09	Wiley Ubuntu Linux Secrets Apr 2009 eBook-BBL in Ebooks
09-juil-09	Public Speaking Workbook - in Ebooks
09-juil-09	10 Tips For Successful Public Speaking - in Ebooks
09-juil-09	The JavaScript Anthology 101 Essential Tips Tricks and Hacks - in Ebooks
10-juil-09	Bit Torrent For Dummies in Ebooks
09-juil-09	CISSP - Certified Information Systems Security Professional Study Guide, 2nd Edition (2004) - in Ebooks
10-juil-09	The Fat Burning Bible in Ebooks
10-juil-09	Wrox Professional ASP NET 3 5 SP1 Edition In C Sharp and VB Apr 2009 eBook-BBL in Ebooks
10-juil-09	Dynamic Learning Adobe Dreamweaver CS3 in Ebooks
09-juil-09	The Object-Oriented Thought Process (3rd Edition) -ebooks- in Ebooks
09-juil-09	Statistics in a Nutshell A Desktop Quick Reference -ebooks- in Ebooks
09-juil-09	Statistics For Dummies -ebooks- in Ebooks

Recherche au 10/07/09 dans la catégorie « ebooks »

Annexe 3 : 50 résultats les plus récents d'une recherche sur Mininova (torrent)

Date d'ajout	Description du fichier
16-juin-09	histoire empire ottoman 18 livres in Ebooks
29 May 09	Clive Cussler Collecton 16 Livres in Ebooks
22 May 09	(Livre Audio) Le Petit Poucet, de Charles Perrault in Audio books
22 May 09	(Livre Audio) La Belle Au Bois Dormant, de Charles Perrault in Audio books
21 May 09	(Livre Audio) Cendrillon Ou La Pantoufle de Vair, de Charles Perrault in Audio books
21 May 09	livre ovnis l 'armée démasquée [french] in Articles
15 May 09	Gaston-Leroux---L'Homme-Qui-A-Vu-Le-Diable-(Livre-Audio) in Audio books
14 May 09	Edgar-Poe---Le-Portrait-Ovale-(Audio-Livre) in Audio books
26 Apr 09	LIVRE-AUDIO-DU-26-04-2009 in Audio books
24 Apr 09	PACK-LIVRE-AUDIO-DU-24-04-2009 in Audio books
23 Apr 09	cod de la route 2009 le livre in Manuals
22 Apr 09	livres anticonformistes in Essays
21 Apr 09	LIVRES-AUDIO-DU-21-04-2009 in Audio books
19 Apr 09	LIVRES-AUDIO-DU-19-04-2009 in Audio books
19 Apr 09	Allan Kardec - Le Livre des mediums (Livre Audio) in Audio books
16 Apr 09	PACK-LIVRES-AUDIO-DU-16-04-2009 in Audio books
15 Apr 09	LIVRES-AUDIO-DU-15-04-2009 JACQUES BREL RACONTE in Audio books
14 Apr 09	LIVRES-AUDIO-DU-14-04-2009 in Audio books
14 Apr 09	Apostila Livre de Reg ncia para Orquestras - CCB in Articles
13 Apr 09	PACK-LIVRE-AUDIO-DU -13-04-2009 in Audio books
29-mars-09	Harry Potter---Livres audio 1 a 4 by Damo CHAMPLEMOUNIER in Audio books
10 Feb 09	[CHEPCHEP]Nephilim 3e - Livre du joueur (JdR) in Ebooks
10 Feb 09	[CHEPCHEP]Nephilim 3e - Livre du joueur (JdR) in Ebooks
09 Feb 09	harry potter la suite complete livre pdf ebooks in Ebooks
08 Feb 09	livre cuisine en image de cap cuisine in Ebooks
08 Feb 09	livre de cuisine en image tres bien fait in Ebooks
05 Feb 09	livre de cuisine en image cap cuisine in Ebooks
05 Feb 09	LIVRE CUISINE EN IMAGE Cap Cuisine in Ebooks
27-janv-09	5 Livres de Médecine Générale (2) [french] in Articles
26-janv-09	5 Livres de Médecine Générale (1) [french] in Articles

20-janv-09	Michael Connelly 1993 Les Egouts de Los Angeles Livre Audio in Audio books
14-janv-09	BD - Les Aigles de Rome - Livre I [french] in Articles
13-janv-09	Livre de bricolage pour enfants) [french] in Articles
12-janv-09	Le Grand Livre de Cuisine d'Alain Ducasse in Ebooks
10-janv-09	Grand Livre de Cuisine d'Alain Ducasse [Desserts et Patisseries] in Ebooks
07-janv-09	livre de massage [french] in Articles
30 Dec 08	Sherlock Holmes - 12 Livres audio de Sir Arthur Conan Doyle [french] in Audio books
29 Dec 08	Grand Livre De Cuisine D Alain Ducasse [FRENCH] in Ebooks
22 Dec 08	La Religion Islamique[AIO 255in1][articles-audios-fatwas-livres-videos] in Ebooks
20 Dec 08	Livres Islamiques in Ebooks
20 Dec 08	Le livre Photoshop CS des photographes numeriques FR [french] in Articles
28-nov-08	Coupable D 'être Innocent Jacques Mesrine (livre) [french] in Articles
12-nov-08	RACONTES MOI DES HISTOIRES 26 LIVRETS [french] in Comics
11-nov-08	(MT) BILAL livres d art in Ebooks
24-oct-08	Farcry2 livret pdf in Manuals
22-sept-08	Michel Zévaco LES PARDAILLAN - Livre 2: L'Épopée d'Amour fr in Audio books
22-sept-08	Livres sur la programmation in Ebooks 4 commentsMichel Zévaco LES PARDAILLAN - Livre 1: Le Chevalier de Pardailan fr in
19-sept-08	Audio books
25-juil-08	Grand-Livre-de-Cuisine-Alain-Ducasse-Mediterranee pdf in Ebooks
25-juil-08	Grand-Livre-de-Cuisine-Alain-Ducasse-Desserts-Patisseries pdf in Ebooks

Recherche en date du 17/06/09 sur le mot-clé « livre » dans la rubrique « ebooks »

Annexe 4 : Top 50 des sources de livrels sur Demonoid (torrent)

Titre du fichier	Genre	Téléchargements	Seeders	Leechers
Larousse gastronomique	pratique	3847	2	1
Harry Potter : tomes 1 à 7	roman	2057	14	3
Bescherelle : compilation	scolaire	1699	7	0
1300 pièges du français	scolaire	1574	8	0
Assimil : l'espagnol sans peine	pratique	1079	7	1
Assimil : l'arabe sans peine	pratique	760	5	1
Histoire de France en BD : tome 1	BD	672	2	0
Derrida : livres	essai	587	1	3
Le Petit Prince	roman	545	3	0
Assimil : l'italien sans peine	pratique	540	2	1
Assimil : l'anglais américain	pratique	537	2	0
Arthur Rimbaud : œuvres complètes	poésie	428	2	0
Le nouveau russe sans peine	pratique	409	0	0
La langue gauloise	scolaire	403	2	0
Cours de mathématiques MP	scolaire	295	2	1
Potagers de poche	pratique	270	2	0
Le Verger tropical	pratique	270	0	0
Cours de linguistique générale	essai	262	5	0
Les Fourmis	roman	195	0	0
Cycle de Fondation	roman	193	4	0
Jardiner avec le changement climatique	pratique	184	3	0
La Vision intégrale	essai	179	6	0
Assimil : le polonais sans peine	pratique	173	1	0
Donjons & Dragons 4e édition	JdR*	169	3	1
Rêve de dragon	JdR *	140	1	0

Les Bidochons : tomes 1 à 9	BD	135	4	0
Twilight : tome 2	roman	132	4	0
Appel de Cthulhu : Delta Green	JdR *	131	3	0
Twilight : tome 3	roman	118	5	0
1984	roman	112	6	0
Emergences résurgences	poésie	109	4	0
Twilight : tome 1	roman	108	4	0
Appel de Cthulhu 5e édition	JdR *	105	0	0
Cycle d'Ender : tome 1	roman	101	1	0
Zazie dans le métro	roman	97	2	0
Le Second Fils de l'homme	roman	96	1	0
Hawkmoon 2e édition	JdR *	94	0	0
Elric	JdR *	91	0	0
JRTM	JdR *	84	1	1
Les Bienveillantes	roman	76	1	0
Bloodlust : Livre des règles	JdR *	74	4	0
Unknown Armies	JdR *	62	4	1
James Joyce, l'Irlande, le Québec, les mots	essai	53	4	0
Louis XIV et le Siam	essai	31	3	1
Saint Seya : Hades 01	BD	28	4	1
L'Etat voyou	essai	27	4	0
Le Mystère de l'évolution	essai	6	5	1
Le Maître de zen	roman	3	3	0
Rien ne va plus en physique	essai	3	2	0
Histoire de l'anesthésie	essai	1	2	0

*Recherche en date du 17/07/09 dans la catégorie « Books » en français, filtrée pour exclure la presse.
« Téléchargements » indique le nombre de téléchargements complets du fichier, tels qu'indiqués sur Demonoid.
* Jeu de Rôles*

Annexe 5 : Top 50 des partages de livrels sur eMule

Nom du fichier	Type
Pierre Desproges : collection 12 CD	livre audio
Clara Morgane : Le livre du Kama Sutra	beau-livre
Grand livre de cuisine d'Alain Ducasse	pratique
Albert Camus : L'Etranger	roman
Le Jardin des insectes	pratique
Marc Levy : la Prochaine fois	livre audio
Stieg Larsson : trilogie complète	roman
Bernard Werber : Encyclopédie du savoir relatif et absolu	essai
Recettes méditerranéennes	pratique
L'Electricité	pratique
Le Livre de Satan	BD
Mon grand livre de projets créatifs super amusants	pratique
Pierre Desproges : dictionnaire superflu	essai
Saint Exupéry : le Petit Prince	livre audio
Les hommes viennent de Mars et les femmes de Vénus	livre audio
Assimil : le japonais sans peine	pratique
Ernesto Che Guevara : Voyage à motocyclette	roman
Fred Vargas : l'Homme aux cercles bleus	livre audio
Bernard Werber : les Fourmis	livre audio
J.K. Rowling : Harry Potter et la chambre des secrets	roman

Jonathan Littell : Les Bienveillantes	roman
Bret Easton Ellis : American Psycho	roman
Sœur Marie-Thérèse des Batignolles	BD
Les 50 meilleures recettes de quiches	pratique
Stephenie Meyer : Twilight, tome 1	roman
JC Grangé : Les rivières pourpres	livre audio
Tartes et salades de Sophie	pratique
Comprendre les mathématiques CE1	scolaire
In Nomine Satanis 2e édition	JDR
J.K. Rowling : Harry Potter et l'ordre du Phénix	roman
Nicolas Hulot : le petit livre vert	essai
Bernard Werber : les Thanatonautes	roman
Umberto Eco : le Pendule de Foucault	livre audio
Manuel pratique de cuisine provençale	pratique
Histoire secrète de l'ordre des Templiers	BD
La Cuisine niçoise d'Hélène Barale	pratique
Assimil : pour mieux connaître l'arabe	pratique
Astérix et la grande traversée	BD
Assimil : le brésilien sans peine	pratique
Bouteflika, une imposture algérienne	essai
Hugdebert : les Reines rouges	BD
Paulo Coelho : l'Alchimiste	roman
Adolf Hitler : Mein Kampf	essai
Stephenie Meyer : Twilight, tome 2	roman
Stephenie Meyer : Twilight, tome 3	roman
Stephenie Meyer : Twilight, tome 4	roman
Astérix en Corse	BD
La Programmation orientée objet en PHP	pratique
Sigmund Freud : 5 leçons sur la psychanalyse	essai
Aldous Huxley : le Meilleur des mondes	roman

*Recherche du 18/07/09, cumulée sur les mots-clés : « ebook french », « ebook fr », « livre », sur 4 serveurs différents.
Résultats filtrés, limités aux livres sous droits, et classés par nombre de seeds*.*

Remerciements

L'auteur souhaite remercier quelques personnes qui ont contribué, de près ou de loin, à l'accomplissement de l'étude : Rémi Douine, Hubert Guillaud et Constance Krebs pour leur relecture attentive, ainsi que Brian O'Leary (Magellan Media), Paul Léveillé et Guénaël Amieux.

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.