

EbookZ?

Study of the availability of illegal downloads of books in French on the Net in 2009

Mathias Daval (Edysseus Consulting) | LE MOTif | October 2009

Complete study downloadable at: www.lemotif.fr

Illegal downloading is an important aspect of the general issues of copyright protection and the distribution of income among the various book industry actors, as well as, even more broadly, the future of the publishing industry in an increasingly digital world.

The economic consequences of digital piracy have not yet been closely studied. Some industry professionals are maintaining a wait-and-see attitude while many others have adopted a more alert stance, but they all share the disadvantage of having to base their reactions on fantasies rather than real foundations because until now there has been no concrete investigation of book piracy on the Net.

Illegal downloads: the first survey

Illegal downloading of books on the Net is not a new phenomenon, but since mid-2008 it seems to have taken a global quantum leap.

There have been very few studies of book piracy on the Net. So far this parallel market has been quite marginal, dwarfed by the piracy of music, movies and video games. Still, if we want to avoid relying on the most fuzzy conjectures regarding the real volume of illegal downloading and the consequent revenue loss, first we have to take a look at what is available on the various pirate networks.

The study initiated by Le MOTif this past summer **was the first in France to specifically analyze the illegal availability on books online**, and to provide answers to the following questions:

- What exactly is available online?
- Which books and publishers are at risk?
- How do these pirated editions circulate?
- What are the real quantities involved?

The subject of this study is “the digital piracy of books” and not “the piracy of digital books.” Since in the English publishing world the term “book” can mean a magazine as well as classical books, when discussing illegal downloading, currently the term **ebook** (plural: “**ebookz**,” sometimes also called “**bookwarez**”) covers not only books but periodicals as well, with the latter constituting 70-80% of the files available for illegal downloading. We have excluded periodicals from our study, which is narrowly focused on **print books and audio books in French published in France**, and, further, concerns only **for-pay and copyright-protected content**.

It would be more accurate if instead of “piracy” we were to speak of the illegal uploading and/or downloading of books in digital format. The term “pirate” in French judicial use more often

refers to “knockoffs” or counterfeit merchandise manufactured for profit, which does not describe the current situation for books, as our research for this study has demonstrated.

It’s not easy being a pirate

Most book piracy today, while sometimes carried out by organized teams, still involves the labor-intensive process of scanning the paper version of a book, a task that takes several hours of painstaking work to achieve quality results. This contrasts strongly with the immediacy of making a digital copy of a movie or music.

Once the file is created, it is posted on distribution circuits. The main ones are **direct downloads, ed2K and BitTorrent**. It can then be downloaded by Web users using procedures of varying degrees of complexity depending on the circuits and titles involved.

Methodology

To evaluate the availability of “pirated” content we explored the main distribution circuits (those mentioned above as well as newsgroups, the Internet Relay Chat and HTTP). Based on these observations we established a series of ensembles or samples, focusing on the most easily available files.

The files were organized into three categories: **print books, comic books and graphic novels, and audio books**. These are the ensembles we have analyzed, with an overview of the results offered here. (The complete study can be downloaded at: www.lemotif.fr).

Less than 1 percent of print-edition titles

An evaluation of the number of books available for illegal downloading remains complex. Nevertheless we estimate that during the summer of 2009 the total was **between 4,000 and 6,000 titles**, of which 3,000-4,500 were comic books and graphic novels. This number represents **less than 1% of the titles legally available in print format, which is still a very small amount**.

After three months of research covering a very large number of illegal files posted on the Net on a variety of distribution circuits, the **following are our estimates** for titles in French:

- ❖ Approximately **1,000-1,500 e-book titles**: a large part of these books are scientific, technical or medical e-books.
- ❖ Approximately **3,000-4,500 comic book and graphic novel titles**: this considerable number of titles is largely due to the distribution of complete series (for example, the collected *Lucky Luke* books).
- ❖ Approximately **200-300 audio book titles, of which at least half** are in the public domain.

	Pirated titles	Legally available titles	% pirated titles
Books and graphic books	4,000–6,000	565,000	0.7% - 1.1%

Heavy files

The **average size of a pirated book is 29.7 Mo**. This average is so heavy because of the large number of illustrated books available for illegal downloading, many of them textbooks and practical titles. The most common format is **PDF**.

	PDF image	PDF text	Word	TXT	DjVu	HTML	IMG
Total	38.6%	29.4%	4.8%	0,7%	6,2%	0.2%	20.1%
Books	38%	42%	6.8%	0.8%	8.8%	0.3%	3.3%
Graphic books	36.1%	0%	0%	0%	0%	0%	62.9%

File format for books and graphic books, excluding audio books.

The top 20 pirated authors

Gilles Deleuze, Bernard Werber and Amélie Nothomb are the authors with the most titles available for illegal downloading.

	Author	Pirated titles
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	3
19	Stephenie Meyer	3
20	Ray Bradbury	3

Notable points:

— **As expected, a great many authors of bestsellers:** Half of the names on this list are among today's best-selling authors in France (Bernard Werber, Amélie Nothomb, Frédéric Beigbeder, J.K. Rowling, Michael Connelly, Daniel Pennac, Marc Levy, Paulo Coelho and Stephenie Meyer). Their strong presence on this list is the logical correlation of their popularity with the general public.

— **As well as works by certain categories of writers:**

• **Philosophers:**¹ 25% of the top 20 authors are philosophers (which is not the case with print best-sellers), and one of them is first on the list. There are several reasons for this, including the reputation these writers enjoy abroad (where Net users are often eager to avoid paying the high cost of imported titles that are not always available in local bookstores or libraries), and, more generally, the intense Web usage among students and academics. Therefore members of the general public are far from the only people to resort to illegal downloading.

• **Science fiction and fantasy writers:**² they also represent 25% of the list. This doubtlessly reflects the cultural tastes of the average Net downloader.

— **And esoteric authors:** among them Eckhart Tolle and Ken Wilber, two authors often categorized as “religious and esoteric,” a grouping overrepresented on the illegal networks in comparison to its legal distribution in print.

The top 20 pirated titles

The titles most often listed and shared on the various networks are *Le Sexe pour les nuls*, the *Harry Potter* series by J.K. Rowling, Alain Ducasse’s *Grand Livre de cuisine* the *Twilight* series by Stephenie Meyer.

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

* Total of the various volumes in the series / ** Also on the list of the authors with the most pirated titles

The publishers at risk

The three French publishers with the most pirated titles are **Gallimard**, **Dunod** and **Hachette**. Regarding graphic books, they are **Delcourt**, **Dargaud** and **Dupuis**.

¹ Gilles Deleuze, Jean-Paul Sartre, Albert Camus, Michel Foucault, Paul Ricoeur.

² Bernard Werber, J.K. Rowling, Isaac Asimov, Stephenie Meyer, Ray Bradbury.

An overrepresentation of non-fiction and practical books.

Each of the three categories “**Non-fiction**,” “**novels**” and “**practical books**” represent a little more than 25% of pirated titles. Here non-fiction and how-to books are overrepresented in comparison with their print edition sales.

	Illegal files	Legal print edition sales*
Essays and documents	28.8%	8.7%
Novels	27.1%	25.6%
Practical	26.7%	10.8%
School & home study	12.9%	12.1%
Children’s books	2.2%	17.4%
Coffee-table books	0.9%	2%
Poetry & theater	1.4%	0.6%

**Percentage of copies sold (Source: SNE, 2009)*

Regarding the subjects we note a concentration of **STM, philosophy, cooking, science fiction and fantasy**, and **computer** titles.

The currency of print titles has little impact on their online availability

Piracy usually involves **contemporary titles**: two out of three pirated titles were published within the last ten years, but the **latest publications are not those most often pirated**: only one in four pirated books was published within the last four years.

	1900-1950	1951-1970	1971-1990	1991-2000	2001-2005	2006-2007	2008-2009
Books	1.1%	4%	10.5%	21.9%	36.4%	18.2%	7.9%
Graphic books	0%	0.2%	4.1%	32.8%	29.8%	17.4%	15.7%

Only **8%** of the books on the **top 50 list of bookstore print edition sales** (*Livres-Hebdo*, 2008) are available for illegal downloading.

The unavailability of certain titles in a legal print edition is only one motivation for illegal downloading and not in itself determinate: **three out of four pirated titles are available in print**.

Looking at it from the opposite angle, **25.6% of the books and 31.4% of the graphic books are no longer available in legal print editions**. These percentages reflect that fact that one of the motivations for illegal downloading is to have access to titles that are out of print or unavailable.

They demonstrate **the importance for publishers of developing the most exhaustive possible legally available current catalogue**. Nevertheless, since the overwhelming majority of pirated titles are available for legal purchase, there are other motivations also driving illegal

downloading. One is cost,³ but the most important is that some people specifically seek access to content in digital form.

In this regard, it must be added that among the files we analyzed, **94.9% of the pirated titles are not legally available in** digital form.

Conclusion

The extent of piracy should neither be exaggerated nor ignored. We would like to offer publishers the following recommendations:

>> Expand their digital catalogues. This is the key condition for being able to compete with illegal downloads. Further, they should make available high-quality legal e-books (most pirate files are also high-quality) at attractive prices.

>> Step up their surveillance of the online distribution of their titles and their alert systems. The purpose is not only to regulate downloading when necessary, but above all to better understand the demand for their titles and their attractiveness in the eyes of the public.

>> Work with governmental bodies to set up, as soon as possible, professional training programs designed for independent publishers so as to avoid the creation of a digital gap between major publishers and small publishing houses.

This is an **initial survey** of illegal online availability, a sort of **snapshot of an inherently volatile phenomenon**. It should be supplemented by reference to earlier research, especially that focused on the users of illegal downloading, as well as **new and regularly repeated observations**.

Further, any real consideration of the questions of “piracy” and digital publishing must take into account overarching issues such as the use value of books, the role of authors in the book industry, and the best way to ensure that their rights are respected and that they are properly remunerated.

³ The average purchase price for a legal paperback edition of a pirated title in all the categories taken together is 15.70 € for books and 12.40 € for graphic novels. This is close to the average price for a paperback in general.