



HAL
open science

The Digital Strategy of Editorial Houses: A Matter of Book Content?

Elisa Salvador, Pierre-Jean Benghozi

► **To cite this version:**

Elisa Salvador, Pierre-Jean Benghozi. The Digital Strategy of Editorial Houses: A Matter of Book Content?. International Journal of Arts Management, inPress. hal-02962327

HAL Id: hal-02962327

<https://hal.science/hal-02962327v1>

Submitted on 9 Oct 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

This article has been accepted for publication by the *International Journal of Arts Management*. It can be cited as follows:

Salvador E., Benghozi P.-J., “The Digital Strategy of Editorial Houses: A Matter of Book Content?”, *International Journal of Arts Management*, forthcoming 2021.

‘The Digital Strategy of Editorial Houses: A Matter of Book Content?’

« La stratégie numérique des maisons d’édition : une question de genre de livres ? »

*** Elisa Salvador, ** Pierre-Jean Benghozi**

*** Elisa Salvador: ESSCA School of Management, Paris, France.**
corresponding author, e-mail: Elisa.SALVADOR@essca.fr, phone +33141862001

**** Pierre-Jean Benghozi:** Research Director CNRS and Professor at the **Ecole Polytechnique, Paris, France**, (Management Research Center, I3-CRG, Ecole Polytechnique, CNRS, Université Paris-Saclay, Bt Ensta 828, Boulevard des Maréchaux 91762 Palaiseau Cedex, France), and Professor at **Geneva University (Geneva School of Economics and Management)**, e-mail: pierre-jean.benghozi@polytechnique.edu

Keywords: cultural and creative industries; book publishing industry; e-book; e-readers; book catalogue

Elisa Salvador holds a higher doctorate (Habilitation à Diriger des Recherches, HDR) in economics sciences and management conferred from Paris13 University (October 2019) and an international PhD in Institutions, Economics, & Law from the University of Turin (Italy). She has worked for the Italian National Research Council (CNR) on several projects focused on innovation policies. She won the CNR award “Promotion of Research 2005” for her project “The financing of research spin-offs: An analysis of the Italian case”. She has collaborated with the Polytechnic of Turin and with the ESCP-Europe Business School and she has been adjunct Professor at Iéseg School of Management, Economics and International Business Economics tracks (2014-2016). She has been a researcher at Ecole Polytechnique, Paris (2012-2015) working on a project focused on R&D and innovation in cultural and creative industries. She is currently permanent professor at ESSCA School of Management, where she is the coordinator of the Master course “Managing Creativity and Innovation”.

Pierre-Jean Benghozi is Research Director at the National Centre for Scientific Research (CNRS) and professor at Ecole polytechnique (Paris) and at University of Geneva (GSEM). Since the early eighties he has developed a pioneering research group on Information Technology, Telecommunications, Media and Culture. Co-Chairman of AIMAC, the largest international Conference for Art and Culture Management, he is also a board member of scientific committees in the highest French institutions, international scientific conferences and academic journals, as well as former commissioner of the executive board of the French national regulatory authority for electronic communications (ARCEP).

A preliminary but significantly different version of this paper was presented at the 13th International Conference on Arts & Cultural Management (AIMAC), which was held on June 26 – July 1, 2015, in Aix-Marseille, France. It has also been presented at an internal seminar at Ingenio (Csic-Upv) Universitat Politècnica de València, Spain, on the 5 of October 2017.

ABSTRACT

Traditional editorial houses are now confronted with the emergence of the online diffusion of e-books. This phenomenon has forced publishing houses to also make their catalogue of printed books available as digital versions. However, it would be inaccurate to envision and report on the digital strategies of publishing houses in a global way since digitisation has very different effects according to the various types of books: dictionaries, illustrated books or novels, for example. The aim of this article is therefore to examine to what extent the type and categories of books influence — or not — in a differentiated way the behaviour of publishers in their digitisation strategies and the design of their e-book catalogues. The expected results should reflect divergent approaches towards proactive behaviour and the appropriation of e-book technologies. Not all publishers are on the path towards publishing in digital formats. Consequently, this article provides a comparative analysis of catalogues of printed books and e-books. It compares the strategy of 104 editorial houses, focusing on a specific country – in this case, France. The most widely diffused book categories with the total number of books available in printed and digital formats have been identified for every editorial house analysed. The analysis is supported and processed thanks to the use of the national systematic databank of French publishers (*Electre*). The results shed light on the effective influence of the digital book on printed book categories and make it possible to characterise accordingly the strategies of editorial houses in the digital age.

RESUME

Les maisons d'édition traditionnelles sont aujourd'hui confrontées à l'émergence de la diffusion en ligne de livres électroniques. Ce phénomène a contraint les maisons d'édition à rendre leur catalogue de livres imprimés également disponible en version numérique. Cependant, il ne serait pas juste d'envisager et rendre compte des stratégies numériques des maisons d'édition de manière uniquement globale. Car la numérisation a des effets très différents selon les différents genres de livres : dictionnaires, livres illustrés ou romans, par exemple. L'objectif de cet article est donc d'examiner dans quelle mesure le type et les catégories de livres influencent - ou non - de manière différenciée le comportement des éditeurs dans leurs stratégies de numérisation et la conception de leurs catalogues de livres électroniques. Tous les éditeurs ne sont pas engagés de la même manière pour publier en formats numériques. Notre hypothèse était que l'analyse des catalogues reflète des approches divergentes en matière d'anticipation et d'appropriation des technologies du livre électronique. Cet article propose une analyse comparative des catalogues de livres imprimés et de livres électroniques. Il compare la stratégie de 104 maisons d'édition, en se concentrant sur un pays spécifique, en l'occurrence la France. Nous avons identifié et analysé, pour chacune de ces maisons d'édition les catégories de livres les plus diffusées ainsi que le nombre total de livres disponibles en formats imprimés et numériques. Cette étude s'appuie sur la banque de données systématique nationale des éditeurs français (*Electre*). Au final, les résultats mettent en lumière l'influence effective des catégories de livres imprimés sur la disponibilité des livres numériques. Ils permettent de caractériser en conséquence les stratégies des maisons d'édition à l'ère numérique.

Introduction

Cultural and creative industries (CCIs) have attracted ever-increasing interest in recent years: several works underline the ‘test center’ nature of these industries in relation to other sectors, and more particularly in the face of the disruptions and transformations brought about by digital technologies, since CCIs were the first to be profoundly affected and challenged by information and communication technologies (ICTs).

In this context, the book industry is a particularly interesting case to investigate because of its intrinsic importance as well as what we learn from the distinctive nature of the transformations it is currently undergoing.

The publishing industry is one of the oldest cultural industries in its structure (Howard, 2009). Furthermore, books are currently the third most likely product to be purchased online in Europe (FEP, 2016).

However, disruptive innovations have destabilised the traditional and secular business model of the book publishing industry. The book industry was affected very early on by the development of digital technologies, but in very different ways that have contributed to a lack of clarity in the development of publishers' response strategies. The first was the development of the online sale of physical books, with the rise of dominant platforms such as Amazon. This was followed by the dematerialisation of books and finally the rise of e-books, driven in particular by the explosion of smartphones and e-readers. Finally, the development of self-publishing and self-editing, made possible by digital technologies that are easy to appropriate, has also contributed to calling into question the visibility of books and the role of publishers and booksellers alike.

Traditional editorial houses are thus confronted with the emergence of the online diffusion of e-books (Benhamou, 2014; Benghozi, Salvador, 2016, 2015, 2014; OECD, 2012; Lebert, 2009). This phenomenon has forced publishing houses to make their catalogue of printed books available as digital versions.

However, it would be inaccurate to envision and report on the digital strategies of publishing houses in a global way since digitisation has very different effects according to the various types of books: dictionaries, illustrated books or novels, for example. The aim of this article is therefore to examine to what extent the type and category of book influence — or not — in a differentiated way the behaviour of publishers in their digitisation strategies and the design of their e-book catalogues.

Currently, because of the size, catalogue or specific nature of their publications, not all publishers are on the path towards publishing in digital formats. The differences in strategies facing the digital age are not only limited to the constitution of catalogues, but they also affect the organisational choices of the publishers themselves.

Consequently, in recent years many newcomers representing small digital editorial houses (i.e. selling only e-books) have appeared on the market, but many traditional book publishers seem to be reluctant to make their complete catalogue available in a digital format.

While new publications are now available in printed and digital formats, according to the authors' copyrights, an editorial house's complete catalogue cannot yet be found in both formats. All book publishers are aware of the importance of dealing not only in printed but also in digital books. The US book market demonstrates that e-books can represent a market share not far from 25% (Rudiger Wischenbart, 2014). Yet, a gap still exists in most countries between print and digital catalogues available, and several reasons are frequently suggested in the literature to explain this: the restrictions raised by digital rights management (DRM), the limitations of publishers' financial capacities for digitising the existing catalogue, their desire to protect the profitability of their current economic model (Simon, 2014) and the lack of skills and mastery of digital technology development (Benghozi, Salvador, 2016). Hence, the revolution brought about by the arrival of the digital book (e-book) has inspired many studies and hypotheses about its influence and impact on the printed book, and even the supposed replacement of the traditional printed book with the e-book (Gilbert, 2015).

Nonetheless, little attention has been paid to the book market (Hjorth-Andersen, 2000) and there are few empirical analyses describing in detail publishers' strategies *vis à vis* the digitisation of their catalogue. The main reasons are both substantive and procedural. On the one hand, except for authors interested only in some specific niche markets (comics, school or scientific publishing), most of the works dealing with the book industry consider the book in general without attempting to compare the different publishing models used by the different categories of books. The other main reason is a methodological one¹. When asked about the size of their printed and digital catalogue, publishing houses generally double the size of their digital catalogue.

¹ As underlined by Waldfoegel, Reimers (2015: 51), ‘Data availability is a major obstacle to the systematic study of the market for books in general and electronic books in particular’.

This happens because a different ISBN² code is assigned to every book and, as a result, the same book available in PDF and E-pub formats is counted twice.

Thus, the book publishing sector has primarily been considered through analyses of the e-book's uses and prices, consumers' willingness, Amazon's role and publishers' strategy in the face of digitisation and book sales, the online retail book market and self-publishing options (cf. among others, Hviid et al., 2017; Simon, 2015; Benhamou, 2015; OECD, 2012; Peltier, Moreau, 2012; Waldfogel, Reimers, 2015; Gilbert, 2015; Latcovich, Smith, 2001), but it has never been considered through a specific focus on catalogues of printed books vs digital formats.

Given this confusing context, this article aims to offer a comparative analysis of the catalogues of printed books and e-books. It deals with the beginnings of the digitisation era (up until the year 2014), highlighting the first strategies followed by the editorial houses. The decision to focus on the way in which the actors elaborate and develop their first responses to a disruptive transformation is particularly interesting in the context where the effects of digital technologies are being felt in all sectors through successive breakdowns, without necessarily building on the experiences or responses developed in other sectors. The main objective of this paper is to scrutinise the hypothesis that publishers embody various digital strategies according to the specificity of their catalogue alongside their different economic characteristics. One could suggest, in particular, that the observable strategy is influenced by the size of the publishers on the one hand, and by the varying content of books (such as text in fiction, pictures and images in comics and information in practical books) and the various types of readers that they address (young readers or heavy readers) on the other. In the first instance, our basic postulate is spontaneously suggested by observations that everyone can easily ascertain. Thus, if the strength of paper media remains undeniable in the reading practices of fiction literature, this is not the case for scientific literature, a growing part of which has directly switched to specialised digital platforms. Similarly, a large part of practical literature (such as cooking or DIY) is also suffering from unprecedented competition from new forms of online publishing.

The aim of this article is therefore to test these hypotheses through an analysis of publishers' strategies in relation to their different categories of books. If our hypotheses are validated, we should find out that the editorial houses who are faced with e-book diffusion have made some book categories more available in digital formats than others, reflecting divergent orientations towards the appropriation of e-book technologies. Differences are also expected to exist according to the size of the publishers: small editorial houses are assumed to be following different strategies from large editorial houses. This analysis is focused on the catalogue 'content' (i.e. category of books) and not on 'sales', and therefore commercial strategies are excluded from the aim of the paper. In fact, Benghozi and Benhamou (2010) emphasise that sales data only partially back up the catalogue structure, namely according to the traditional long tail argument. Rather, online markets have given way to the reorganisation of business models in cultural industries, shifting from the rarity of titles to the technical and operational capacity to choose them, play them and save them in personal spaces; in other words, the emergence of strategic decisions about catalogues of printed books and their digitisation.

Since it is reasonable to assume that the effect of book content on the degree of digitisation is independent of the language or country of the publisher, the choice was made, for methodological reasons of data availability, to focus on only one country – in this case, France – comparing the strategy of 104 editorial houses.

The article is structured as follows: Section 1 presents the main insights about the book publishing industry, with an introduction to e-book diffusion as well as the emergence of e-book platforms; Section 2 focuses on the methodology, while Sections 3-5 present the results of the analysis. Finally, Section 6 discusses the results and provides some concluding remarks as well as issues for future research.

1. The emergence of e-book platforms and e-readers in the book publishing industry

In order to contextualise the effects of digitisation on publishing and to understand how publishers have coped with it, it is important to recall the main features of this industry. The book industry is, like the majority of cultural industries, highly concentrated (De Prato, Simon, 2014). Its structure, commonly described in the economic literature as a fringe oligopoly (Benhamou, 2004), is characterised by a combination of large dominant players together with a myriad of smaller companies, generally positioned in niche or highly

² ISBN (International Standard Book Number) is the normalised product identifier used by publishers, booksellers, retailers and other supply chain participants for ordering, listing, sales records and stock control purposes.

innovative markets, independent or integrated to a greater or lesser extent in consortia. France³ is no exception to the rule. Hachette and Editis dominate the book market just as large publishing groups like Bertelsmann and Holtzbrinck dominate the German market. A large number of medium-sized and small publishers are also present as well as extensive book-selling systems (Barnes & Co., 2012). With regard to ownership structures, the French editorial house Hachette has become a full international player through several acquisitions that have enabled it to establish publishing groups not only in France but also, among others, in the UK, the US, Italy, Spain, Poland, Australia and Canada (Ronning, Slaatta: 2012: 103).

According to Ronning and Slaatta (2012), the ability of niche and specialty publishers to obtain market positions — through their own distribution channels like e-commerce platforms and digital bookstores — provided some opportunity to challenge the oligopolistic trend of this industry while also favouring concentration around the dominant platform Amazon, which benefits from network externalities and its ability to aggregate all segments and niches of the publishing market (Ritala et al., 2014).

The corresponding challenges arise both in terms of international competition and market structure at the national level (due to language specificities) on the one hand, and in terms of the reorganisation of value chains around the emergence of new forms of intermediation, platforms and recommendation mechanisms on the other hand.

In recent years, the traditional book publishing industry has thus been impacted by disruptive and unusual changes (Hviid et al., 2017; Benghozi, Salvador, 2016; Benhamou, 2015; Gilbert, 2015; Lebert, 2009). Books were one of the first products sold on the Internet (Baye et al., 2013), but the e-book, which first appeared on the market in the 2000s, rapidly contributed to the disruption of all the phases of the traditional book value chain. New actors have readily emerged to support innovative tools and new business proposals: the self-publishing option as an alternative to traditional publishing emerged as one of the novelties brought about by digitisation (Hviid et al., 2017). One of the most impressive novelties is the creation of platforms for selling and/or reading (e-)books. These platforms are used to sell books both in digital and printed versions (or in only one of the two versions) and they can compete with a physical bookstore.

The consequence is, of course, that book publishers benefit now from unprecedented opportunities to find, thanks to these new digital platforms, readers through new distribution channels, and to easily address international global markets. On the other hand, they are now obliged, correspondingly, to completely rethink their presence in online markets, first by updating and converting their catalogue (including both the printed and digital offer of books) to make their content available to remote readers, and also by designing their own websites to promote and market their content directly.

In this context, the market structure of publishers has been disrupted by the rise of platforms redefining the value chain around the distribution and logistics functions for certain economic players, and aggregating and referencing product offerings and providing technical interface support for others. Amazon is one of the pioneers in the online distribution of traditional printed books. Since its inception, a plethora of online platforms selling (e-)books has been identified (Benghozi, Salvador, 2014; Baye et al., 2013). These platforms capture new clients as well as their personal data to achieve marketing goals, and consumers are changing their traditional habits: readers buy online instead of in traditional physical bookstores. New intermediaries are emerging, such as software developers offering specific solutions for e-book platforms (Simon et al., 2015).

In the book sector, while Amazon has in recent years reached a dominant position in online book selling, some competing publishers have recently created joint digital platforms for online distribution (Gilbert, 2015; Simon, 2014). Ritala et al. (2014) focused on the ways in which the potential advantages of collaborating with competitors (the so-called ‘coopetition’ strategy) may be obtained through the involvement of competitors in the firms’ business model. To this end, they analysed the case study of Amazon.com’s coopetition-based business models, including in the book sector. The results highlighted that competitors were able to create value for their customers via the Amazon marketplace. This model was particularly beneficial to small bookstores, which had the opportunity to present their offerings to millions of potential customers and to increase the size of the market for e-book sales worldwide. Furthermore, the general ambition of Amazon to create value for customers resulted in the distribution of electronic book content across different and competing platforms (e.g. Amazon.com’s content is available on the Apple iPad through the Kindle app).

The diffusion of platforms is strictly linked to the emergence of e-readers. In this general context, it is worth asking to what extent the emergence of new forms of online distribution has in fact been accompanied by an

³ In general, ‘*The French institutional system favours multi-level governance by virtue of the principle of co-responsibility which enables the different levels to coordinate and strongly support the cultural field*’ (Arnaud, Soldo, 2015: 116 — our translation).

evolution in publishers' publishing strategies and catalogues, or whether the emergence has represented only a marginalist and minimalist impact on the behaviour and editorial practices of publishers vis à vis emerging technologies.

2. Methodology

From a methodological viewpoint, the originality of our empirical investigation lies in the provision of primary data and a detailed analysis undertaken in a given time period (i.e. the first years of digitising catalogues). Given the lack of reliable official statistics to compare catalogues of printed books with those of e-books, this analysis brings primary data taken from the most reliable and recognised French official databank in this field. Our research relied on the construction of a specific database: the list of 104 editorial houses was chosen based on the meaningful size of their catalogues of e-books. The most widely-diffused book categories with the total number of books available in printed and digital formats have been identified for every editorial house analysed.

The case of France is of particular interest because of the market size and the specificities of the role played in general by culture – the so-called '*exception culturelle*' (Lescure, 2013). The French language is also a structuring factor of the book publishing sector and French editorial houses. The pivotal role played by the book industry is proved by the fact that France accounts for 11.9% of the European book market value, in third position just after the UK (12.2%) and Germany (29%), and ahead of Italy (9.3%) and Spain (7.6%) (MarketLine, 2014).

But the prominent role of the French market is not the main reason. It is also interesting because e-book diffusion is increasing,⁴ even though it still comes up against significant delays and the reticence of most publishers, as is the case in many European countries (De Prato, Simon, 2014). In any case, the weak success of the e-book in France compared, for example, to the US market, underlines the specificity of a market more strongly anchored to the printed book. This makes the analysis of the strategic response of editorial houses even more interesting. Last but not least, a focus on the French context is particularly relevant because this is an autonomous market due to the distinctiveness of the language and the specificity of a very protective law on the distribution of books⁵ (Poirel, 2015; Rouet, 2007), meant to preserve editorial diversity by maintaining a dense and decentralised network of bookshops.

Therefore, we have put in place an empirical analysis based on a comprehensive survey of the catalogues of printed books and e-books of 104 editorial houses. Data were collected in June 2014, extracted from the national systematic databank of French publishers, the *Electre* databank (see Annex 1). More specifically, the unprecedented structuring of data and compilation of the catalogues of all editorial houses were achieved as follows.

The editorial houses were selected on the basis of the size of their catalogue of e-books: we picked those with a minimum of 89⁶ e-books. In order to avoid problems of duplicity in counting for the overall digital catalogue, we decided to make a distinction between the various existing digital formats: PDF, Epub and streaming flow.⁷ Only the biggest in size of these formats in the catalogue was analysed for every editorial house, because usually a book is available in PDF as well as in Epub format, for instance. The Streaming format was included in the list because many publishing houses in the comic book category deal with books exclusively in this format.

As well as the size of the general catalogue, we also opted for a distinction between the specific categories of books. We chose two different methods to select and retain relevant categories. Initially, we picked a first set of variables – those regarding the formal characteristics of the books — directly from the *Electre* listing: pocket size (only for the printed catalogue), youth literature, practical books including tourist guides, school books and comics. These variables were refined with a second set, focusing more on the type of book content, extracted from the formal *Dewey* classification (see Annex 2).

Table 1 provides the full list of book categories (i.e. variables for the statistical elaborations) and a summary of the meaning of each category.

⁴ In 2015, the digital book publishing market generated a total turnover of 163.8 million euros, up by 1.5%. This represents 6.5% of publishers' book sales (SNE, 2016: 14).

⁵ The so-called 'Lang Act' establishes that the publisher has the prerogative to set the sale price.

⁶ This solution has been adopted because we substantiated that under the threshold of 89-90 e-books, many data were missing and the leap was significant (i.e. catalogues with only 20-30 e-books on average).

⁷ In fact, some editorial houses consider the same book to be a different item according to its digital format.

Table 1: List of variables identified

<i>Book categories</i>	<i>Content</i>
Printed books	Full catalogue of printed books
Pocket size	Full catalogue of printed books in pocket format
PDF e-book	Full catalogue of e-books in PDF format
Epub e-book	Full catalogue of e-books in Epub format
Streaming e-book	Full catalogue of e-books in streaming format
Youth literature	Full catalogue of literature for the younger generation
Practical books	Full catalogue of practical books and tourist guides
School books	Full catalogue of books for school and extra-curricular books
Comics	Full catalogue of comics for young people and adults
Dewey 000	Computer science, information & general works: this is the most general class, and is used for works not limited to any one specific discipline, e.g. encyclopaedias, newspapers and general periodicals. This class is also used for certain specialised disciplines that deal with knowledge and information, e.g. computer science, library and information science, and journalism.
Dewey 100	Philosophy, parapsychology and occultism, and psychology
Dewey 200	Religion
Dewey 300	Social sciences (sociology, anthropology, statistics, political science, economics, law, public administration, social problems and services, education, commerce, communications, transportation and cultural customs)
Dewey 400	Language, linguistics and specific languages. Literature, which is arranged by language, can be found in Dewey 800
Dewey 500	Pure Science and mathematics
Dewey 600	Technology
Dewey 700	Arts & recreation: arts in general, fine and decorative arts, music, and the performing arts. Recreation, including sports and games, is also classed in Dewey 700
Dewey 800	Literature, including rhetoric, prose, poetry and drama. Folk literature is classed under customs in Dewey 300
Dewey 900	History & geography: the history of a specific subject is classed alongside the subject itself.

In order to investigate the publishers' current catalogues and elicit their various kinds of strategy when faced with the diffusion of e-books, principal component analyses (PCAs) and cluster analyses were used. PCA condenses a large number of variables into a smaller group, called 'components', which reflects most of the possible variability (information) of the original variables. This way, thematic groups were identified by reducing the number of dimensions. The cluster analysis approach makes possible the identification of groups of objects with small within-cluster variations for distinguishing variables and high variations between clusters. The combination of the two methods (Meng, Chen, 2008) turned out to be the best solution for highlighting the main groups and strategic moves of editorial houses. After the identification of five main components contributing to the differentiation of the editorial houses' strategies, the various homogeneous groups of editorial houses emerged by means of the cluster analysis methods (Everitt et al., 2001). The two-step cluster analysis technique was applied for data analysis.

3. Analysis of results: a first outlook

This section focuses on the general description of the catalogues of all 104 editorial houses followed by a detailed drawing of the content of the catalogues of the 100 publishers (see below) finally retained for the PCAs and cluster analyses.

The description of the minimum and maximum sizes of the printed catalogues and the different e-book formats is presented in Table 2. It reveals the prevalence of the pocket version (printed catalogue) and the PDF format (digital catalogue).

Table 2: Descriptive statistics: printed and digital catalogue

<i>Variable</i>	<i>Obs.</i>	<i>Mean</i>	<i>Std. Dev.</i>	<i>Min</i>	<i>Max</i>	<i>Total catalogue 104 ed. houses</i>
Printed books	104	2726.212	6019.142	89	39940	283526
Printed pocket version	104	472.7596	1853.762	0	15709	49167
E-book PDF	104	547.2788	3098.571	0	31509	56917
E-book Epub	104	289.4519	787.5004	0	7064	30103
E-book Streaming	104	143.6346	345.0168	0	1695	14938

The statistical description of the mean, minimum and maximum values of each category (for printed books and for e-books) in the full list of 104 editorial houses displays a huge steady difference between printed and digital formats: the printed value is always superior to the digital one.

While two small publishers were excluded from the analysis because of their very small size and ‘zero’ data in most of the categories of books, L’Harmattan and Gallimard were also left out because they correspond to statistically atypical situations due to the huge difference in size of catalogue compared to the other publishers.⁸ Indeed, these two publishers have a printed catalogue of almost 40,000 books. Nevertheless, these two publishers are likewise very different as regards the size of their various categories in print and digital format: the strategy of the main French publishing houses seems to be very divergent when faced with e-book diffusion. 85% of the L’Harmattan catalogue is available in PDF format, while only 4% of the Gallimard catalogue has been digitised in PDF or ePUB format. Conversely, 39% of the Gallimard catalogue is available in pocket size, while virtually 0% is in pocket size for L’Harmattan. This highlights that Gallimard has a strategy still anchored to the printed catalogue (Black & White), where literature (Dewey 800) makes up 55% of the printed catalogue and 3% of the digital catalogue, and youth literature is 24% of the printed catalogue and 0% of the digital one. L’Harmattan is strongly oriented towards social sciences, or Dewey 300 (40% available printed and 34% in digital formats) and literature, or Dewey 800 (24% available printed and 20% in digital formats).

Considering, more specifically, the variety of book content, the distribution in size of the different Dewey categories (see Annex 2) in the sample of the 100 editorial houses highlights that digital versions are currently poorly represented.

Distribution within the printed catalogue of the last categories collected in the database, i.e. youth literature, practical books, school books and comics, confirms that the publishers with a catalogue in these categories containing over 1,000 books are really in short supply.

Finally, the distribution within the digital catalogue of these same categories (youth literature, practical books, school books and comics) highlights that the publishers with a catalogue of over 1,000 books are few and far between: they are present only for practical books and comics. Youth literature and school books are still very poorly digitised. Regarding the comic book category, as expected, the overall catalogue is available in streaming format.

4. Identifiable structuring basis for catalogue strategies

In order to identify the editorial houses’ specific strategies, we first of all ran a factor analysis. Principal component factor extraction was the method chosen and the Kaiser-Meyer-Olkin (KMO) was the measure of sampling adequacy (Tabachnick, Fidell, 2013). The KMO and Bartlett’s test confirmed the reliability of the PCA’s results. To determine the number of components, the latent root criterion (Eigenvalue > 1.0) was applied: a five-dimensional solution explaining 79% of the variance for the printed catalogue and 81% for the digital catalogue was obtained.

4.1 Components for printed catalogue strategies: the PCA analysis

In order to explore the links between variables and similarities among publishers and to identify the components on which they differ, we performed a PCA on the printed catalogue. The results confirmed our main assumptions and contributed to identifying five main components, each one corresponding to a specific

⁸ These two editorial houses were excluded to avoid a possible influence on the final results and to prevent them distorting the grouping of the cluster analysis. These editorial houses have a catalogue quite disproportionate in size compared to the other publishers for two very different reasons: on the one hand, L’Harmattan develops a very extensive publishing strategy in social sciences and, on the other hand, the Gallimard portfolio benefits from a very old publishing tradition.

publisher's strategic move associated with the categories of printed books. The analysis helped to reduce the rating variables down to five main components, identified as follows (Table 3).

Component 1 is the most comprehensive. It includes the items pertaining to Dewey 100 - Dewey 600⁹ as well as Dewey 800 and Dewey 900¹⁰. Therefore, it can be labelled '*General sciences, literature and technical documentation*'.

Component 2 groups the items pertaining to the youth and comics category; it can be labelled '*Youth and comics*'. Figure 1 shows the pivotal editorial houses associated with the Component 1 and Component 2 strategies.

Component 3 includes items specifically relating to the practical book category, which can be labelled '*Practical documentation*' (Figure 2).

Component 4 groups the items pertaining to computer science (Dewey000); this can be labelled '*Informatics*' (Figure 3).

Component 5 relates to the school book category; it can be labelled '*School publishing*' (Figure 4).

To sum up, the results pointed out that several different categories of books that give rise to and result in different strategies on the part of publishers can indeed be identified. These categories consist of: '*General sciences, literature and technical documentation*', '*Youth and comics*', '*Practical documentation*', '*Informatics*' and '*School publishing*'. Some editorial houses are associated with more than one of these strategic moves: Pearson ERPI (Ed35) mostly characterises the strategy of Components 1, 4 and 5, while Eyrolles (Ed2) is one of the most representative publishers for the strategies of Components 3 and 4. Nouv. éd. de l'université (Ed7) is associated with the strategies of both Components 3 and 5.

Table 3: Printed catalogue: Component Matrix^a

	Component				
	1	2	3	4	5
Pyouth	.249	.793	-.028	.000	.059
Ppractical	.307	.075	.728	-.049	.550
Pschool	.543	-.099	-.523	.256	.528
Pcomics	-.022	.866	-.168	.210	-.045
pDewey000	.286	-.091	.500	.659	-.170
pDewey100	.828	-.153	-.065	-.100	-.258
pDewey200	.619	.074	.031	-.432	-.230
pDewey300	.802	-.168	-.109	-.101	-.314
pDewey400	.705	-.209	-.521	.194	.314
pDewey500	.796	-.097	-.092	.115	-.106
pDewey600	.668	-.132	.355	.457	-.188
pDewey700	.324	.891	-.022	.145	-.039
pDewey800	.466	.226	.080	-.450	-.135
pDewey900	.618	.021	.349	-.400	.404

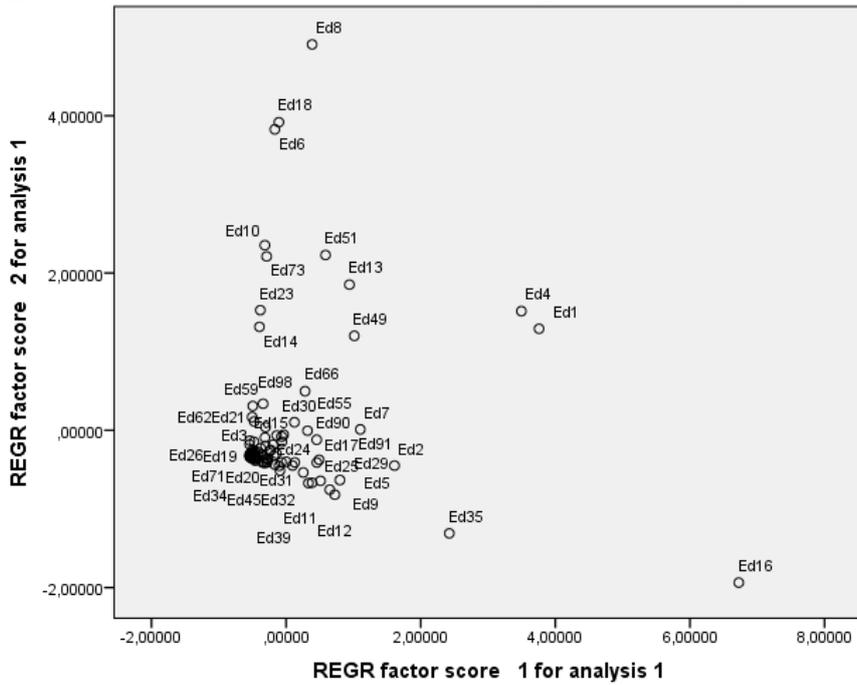
Extraction Method: Principal Component Analysis.

^a 5 components extracted.

⁹ Philosophy, religion, social sciences, language, natural sciences and technology.

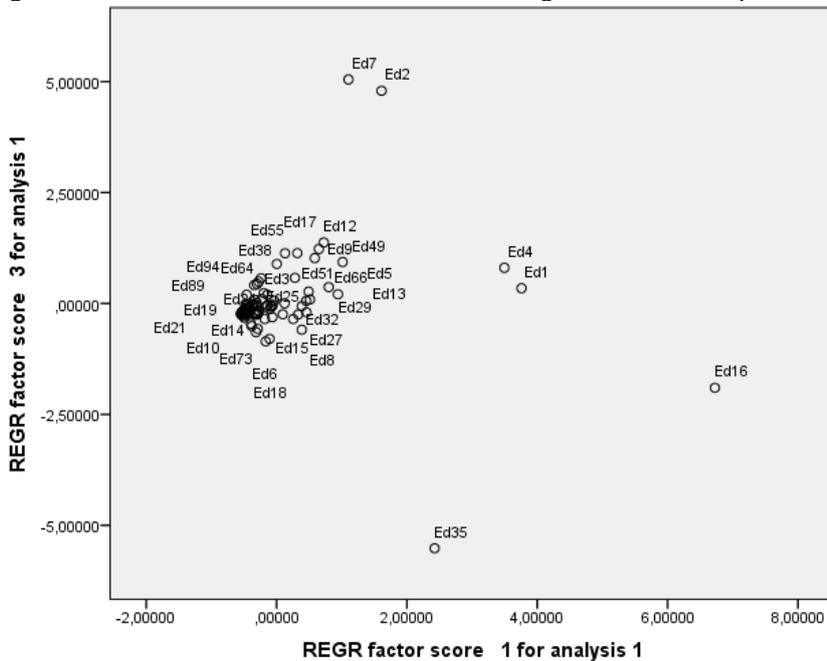
¹⁰ Literature, geography and history.

Figure 1: Editorial houses associated with the strategic moves of Components 1 and 2



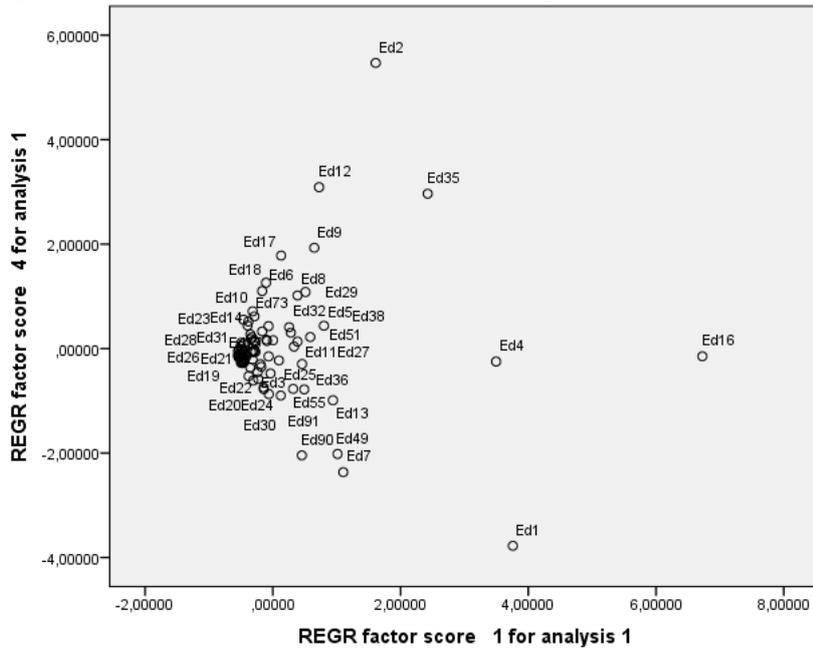
Main editorial houses associated with the Component 1 strategy		Main editorial houses associated with the Component 2 strategy	
Ed4	Flammarion	Ed8	Casterman
Ed1	Seuil	Ed18	Dargaud
Ed35	Pearson ERPI	Ed6	Dupuis
Ed16	PUF		

Figure 2: Editorial houses associated with the strategic moves of Components 1 and 3



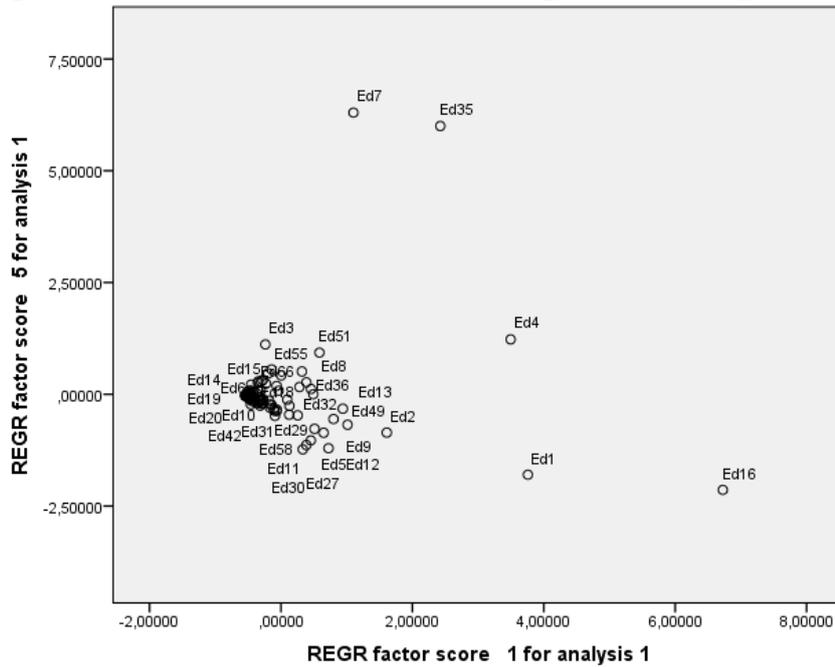
Main editorial houses associated with the Component 3 strategy	
Ed7	Nouv. éd. de l'université
Ed2	Eyrolles

Figure 3: Editorial houses associated with the strategic moves of Components 1 and 4



Main editorial houses associated with the Component 4 strategy	
Ed2	Eyrolles
Ed12	Hermès science publications
Ed35	Pearson ERPI

Figure 4: Editorial houses associated with the strategic moves of Components 1 and 5



Main editorial houses associated with the Component 5 strategy	
Ed7	Nouv. éd. de l'université
Ed35	Pearson ERPI

4.2 Components for digital catalogue strategies: the PCA analysis

Similarly, by extending the same PCA method to the digital catalogues, five main Components emerged, identified as follows (Table 4).

Component 1 is the most comprehensive. It includes the items pertaining to Dewey 000 – Dewey 300¹¹ as well as Dewey 500 and Dewey 600¹². Therefore, it can be labelled ‘*General and social sciences, computer science and technical documentation*’.

Component 2 groups the items pertaining to youth, comics and the Dewey 700 (arts and recreation) category: it can be labelled ‘*Youth, arts and comics*’. Figure 5 shows the pivotal editorial houses associated with the Component 1 and Component 2 strategies.

Component 3 includes items specifically relating to the practical category of books and Dewey 900 (geography and history): it can be labelled ‘*Practical documentation*’ (Figure 6).

Component 4 groups the school and Dewey 400 (language) categories: it can be labelled ‘*School & language editions*’ (Figure 7).

Component 5 relates to the Dewey 800 (literature) category: it can be labelled ‘*General literature*’ (Figure 8).

To sum up, the results from the digital catalogues show that there are just as many differentiated strategies of publishers depending on the type of books they publish. On the other hand, it is interesting to note that, in this case, the categories of books at play do not correspond identically to those observed for printed catalogues. In these cases, the five main publishers’ strategic moves identified regarding the digital catalogue are linked to ‘*General and social sciences, computer science and technical documentation*’, ‘*Youth, arts and comics*’, ‘*Practical documentation*’, ‘*School and language editions*’ and ‘*General literature*’.

Table 4: Digital catalogue: Component Matrix^a

	Component				
	1	2	3	4	5
Dyouth	-.311	.885	-.079	.013	.112
Dpractical	.417	.148	.873	.087	.122
Dschool	.284	.040	-.268	.828	.111
Dcomics	-.339	.924	-.054	.029	.019
dDewey000	.713	.264	.081	-.184	-.373
dDewey100	.790	.219	-.044	-.281	-.081
dDewey200	.616	.130	-.191	-.270	.477
dDewey300	.702	.099	-.217	-.228	.199
dDewey400	.622	.115	-.164	.672	.028
dDewey500	.420	.027	-.237	.302	-.160
dDewey600	.742	.216	.051	-.154	-.446
dDewey700	-.245	.957	.004	.016	.015
dDewey800	.228	-.046	-.303	-.181	.775
dDewey900	.173	.038	.867	.213	.368

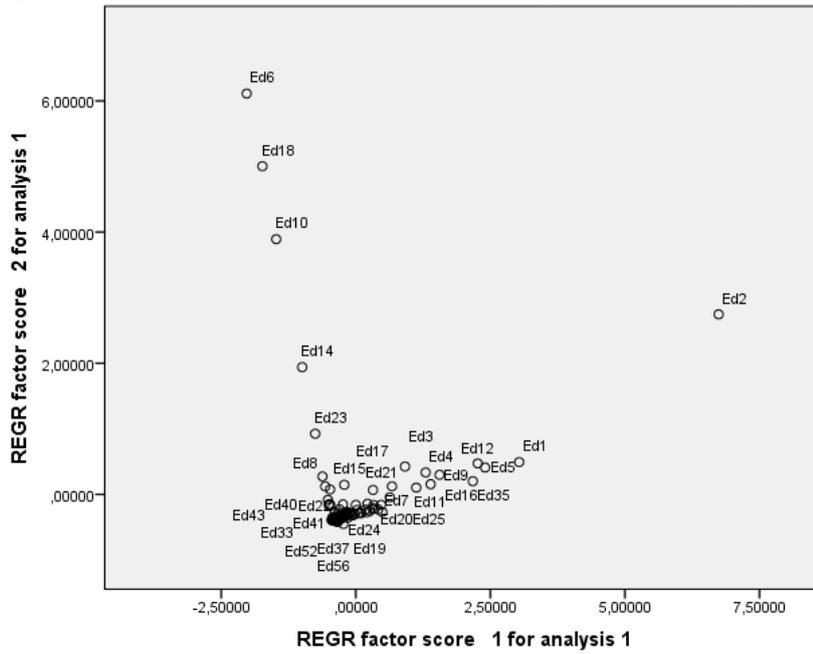
Extraction Method: Principal Component Analysis.

^a 5 components extracted.

¹¹ Computer science, philosophy, religion and social sciences.

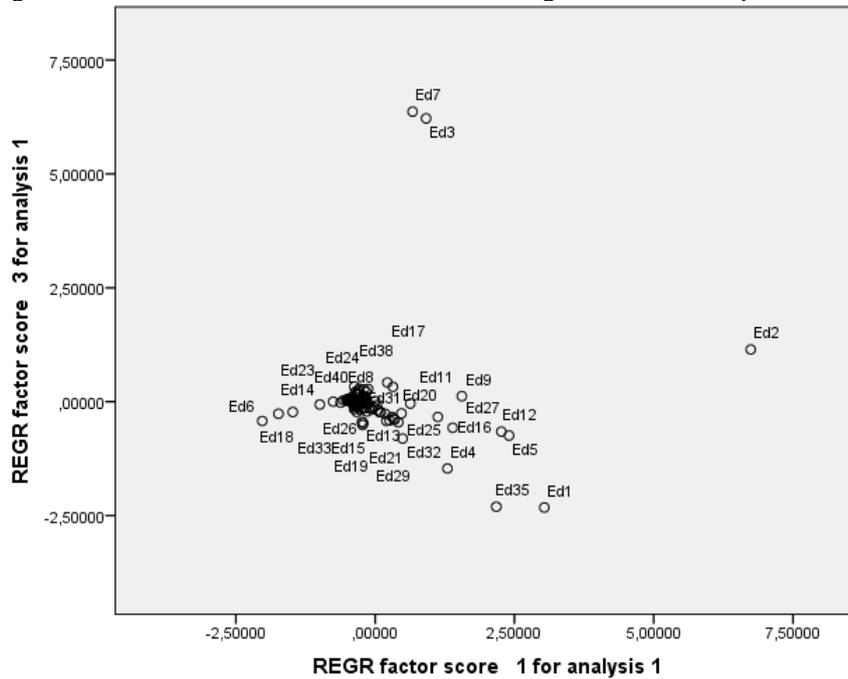
¹² Pure science and technology.

Figure 5: Editorial houses associated with the strategic moves of Components 1 and 2



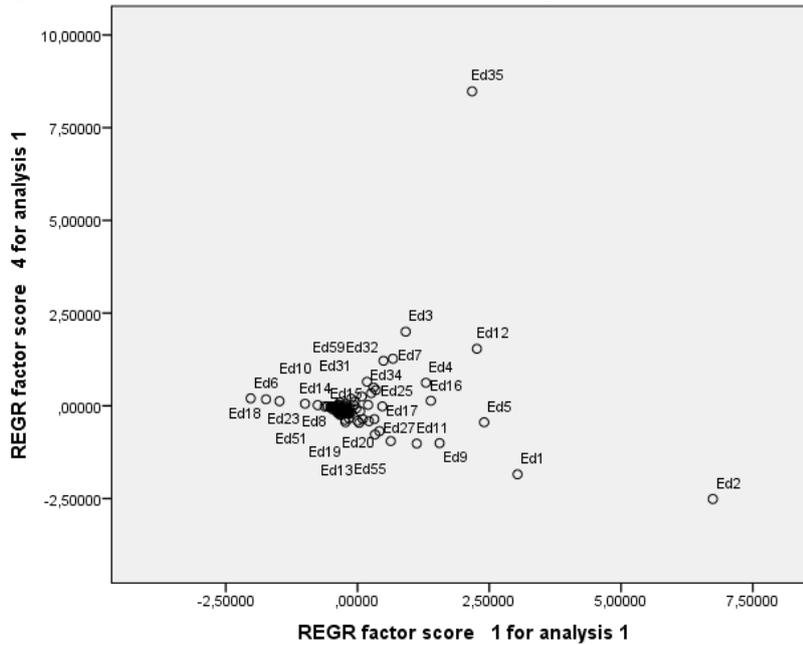
Main editorial houses associated with the Component 1 strategy		Main editorial houses associated with the Component 2 strategy	
Ed2	Eyrolles	Ed6	Dupuis
		Ed18	Dargaud
		Ed10	Le Lombard

Figure 6: Editorial houses associated with the strategic moves of Components 1 and 3



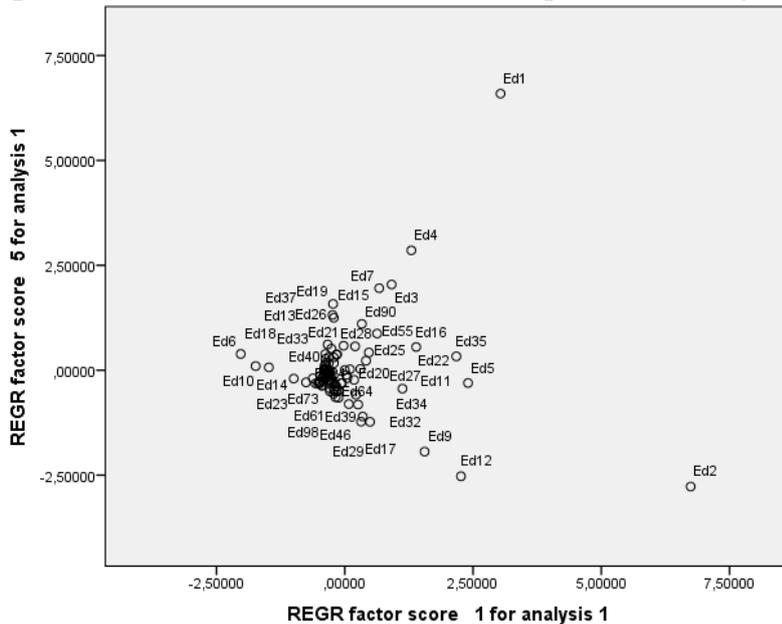
Main editorial houses associated with the Component 3 strategy	
Ed7	Nouv. éd. de l'université
Ed3	Guides de voyage Ulysse

Figure 7: Editorial houses associated with the strategic moves of Components 1 and 4



Main editorial houses associated with the Component 4 strategy	
Ed35	Pearson ERPI

Figure 8: Editorial houses associated with the strategic moves of Components 1 and 5



Main editorial houses associated with the Component 5 strategy	
Ed1	Seuil

5. Publishers' strategies grouping: cluster analysis results

Following the identification and description of the various components and strategic moves identified by the PCA, the next step in the development of the analysis consisted of reducing the heterogeneity of the components and ascertaining the main groups of publishers that share similar strategic moves. To do so, we determinedly ran a cluster analysis. This process enabled us to fine-tune the analysis and ascertain the key

choices of the editorial houses' strategy for their printed and digital catalogues. In fact, the analysis allowed us to characterise sub-groups of publishers for both catalogues (printed and digital).

From a methodological perspective, a satisfactory partition divided the sample into three groups of publishers.

5.1 Publishers' strategies groups for printed catalogues

The application of the two-step clustering, which combined hierarchical and non-hierarchical methods, highlighted three main groups. The Anova table — the classic statistical validation test — confirmed the significance and reliability of these results for printed catalogues. The size of these groups is quite dissimilar: Cluster 1 includes only 4 publishers, Cluster 2 groups 20 publishers and Cluster 3 has 76 editorial houses. Some specific variables characterise certain groups more than others. Dewey 400 (language) is the most important variable for each of the three clusters, followed by the school books (academic) category and Dewey 500 (pure science). Conversely, the youth and practical categories, as well as Dewey 800 (literature) and Dewey 000 (computer science), are not among the most important variables in this clustering.

Consequently, the key features of the three groups can be identified as follows.

Cluster P1: editorial houses with a comprehensive catalogue, including the pocket size format

Cluster P1 is the smallest in size and includes only 4 publishers: Seuil, Flammarion, PUF and Pearson ERPI. The first three publishers are those with a printed catalogue exceeding 15,000 books. Most of the variables strongly characterise this group: the only variables with a low influence are the comics, Dewey 000 (computer science) and practical categories. These features can be explained in part by the fact that Cluster P1 includes three of the main publishers with the largest catalogues of books in 'pocket size' format and literature, or Dewey 800 (Seuil, Flammarion and PUF, while J'ai Lu is included in Cluster P2) and that cover most of the book categories.

Cluster P2: editorial houses with a catalogue more focused on selected categories of books: practical, comics and youth

Cluster P2 includes most of the largest catalogues in the category 'practical books' and it is also well represented in the comics and youth categories, as well as in the pocket size format. It includes 20 publishers. Dewey 000 (computer science) and Dewey 700 (arts and recreation) are also well represented. Among others, this cluster includes the following editorial houses, which are more specialised in youth and comics books: Fleurus, Dargaud, Dupuis, Casterman, Kana, Le Lombard and Bamboo. All of these publishers are among the top publishers in Dewey 700. This Cluster also includes the editorial house Nouvelles Éditions de l'Université that publishes *Le Petit Futé*, one of the most widely distributed travel guide collections in France.

Most of the top publishers for printed and pocket size catalogues as well as for the youth and practical book categories are included in Clusters P1 and P2. Furthermore, as expected, all of the top publishers in the comic book category are included in Cluster P2.

Cluster P3: editorial houses with a more generalist and less specialised catalogue

Cluster P3 is not overly influenced by the pocket size format and is the least influenced by the 'school' book category, as this category is already strongly represented in Cluster P1. This Cluster is the biggest in number (it includes 76 publishers), the most mainstream, and not necessarily specialised in any specific categories, although Dewey 800 (literature) and Dewey 300 (social sciences) are relatively common. CNRS editions, Presses de Sciences Po, Presses polytechniques et universitaires romandes, Lavoisier, Payot and Rivages are some representative publishers included in this Cluster.

To sum up, the cluster analysis clearly identified three main groups of publishers corresponding to specific strategies and catalogue orientation while echoing the global economic dynamics of the publishing industry. Cluster P1 reflects the fringe oligopoly structure (cf Section 1) commonly described in the economic literature, while Clusters P2 and P3 combine the myriad of smaller independent companies, positioned in more or less selected categories of books. According to the above description, we suggest referring to them as '*Black & White*' (Cluster P1), '*Practical & Illustration*' (Cluster P2) and '*Documentation*' (Cluster P3) oriented.

5.2 Publishers' strategies group for digital catalogues

Similar to the printed catalogues, the applied two-step clustering method highlighted three main groups for the digital catalogues, with good significance and reliability of these results according to the Anova table. Also in this case, some specific variables characterise each group more than others. Dewey 700 (arts and recreation) is the most significant variable for each of the three clusters, followed by comics and the youth book category. Dewey 000 (computer science), Dewey 800 (literature) and the school category are not the most important variables in the grouping formation.

Cluster D1: a ‘mirror’ of the printed catalogue Cluster P1, but including more publishers

Cluster D1 (10 publishers) is strongly influenced by most of the Dewey categories as well as by the school category, but not at all by comics. It reflects exactly the same content as Cluster P1 for the printed catalogue. All four publishers in Cluster P1 for the printed catalogue are included here. Nonetheless, Cluster D1 in the digital catalogue is larger than in the printed one, because some publishers that were in Cluster P2 for the printed catalogue are now in Cluster D1 for the digital catalogue (such as Eyrolles, O. Jacob and Ed. d’organisation). All the top publishers in the digital catalogue are included in Clusters D1 and D3.

Cluster D2: the largest and the most mainstream group, analogous to Cluster P3

Cluster D2, the largest group, is hardly influenced by any specific variables except in part by social sciences (Dewey 300) and literature (Dewey 800). Cluster D2 for the digital catalogue is very similar to Cluster P3 for the printed catalogue regarding its size (84 publishers for the digital catalogue vs 76 publishers for the printed one): it includes, among others, all the representative publishers cited in Cluster P3. Again, it is larger because it includes some publishers that were in Cluster P2, such as Actes Sud, J’ai Lu, Fleurus and Novalis.

Cluster D3: the smallest but the most ‘specialised’

Cluster D3 (6 publishers) is the most represented in the comics, practical and youth categories as well as in arts and recreation, or the Dewey 700 (Dupuis, Le Lombard and Dargaud are the three top publishers in Dewey 700 for the digital catalogue) and Dewey 900 (geography and history). The top publishers in the practical books and comics categories are included in this Cluster. Conversely, school books and computer science (Dewey 000), philosophy (Dewey 100), religion (Dewey 200) and literature (Dewey 800) are absent from Cluster D3. The six publishers¹³ included in Cluster D3 here are exactly the same publishers as those included in Cluster P2, with the exception of Guides de voyage Ulysse (included in Cluster P3). So, Cluster D3 for the digital catalogue reflects exactly the content of Cluster P2 for the printed catalogue, but with a different number of publishers (20 for the printed catalogue vs 6 for the digital one) because, as explained above, in the shift from the printed to the digital catalogue, some publishers have been ‘removed’ from Cluster P3 and have re-joined Clusters D1 and D2.

To sum up, it is important to note that the method put in place already confirms some of our initial assumptions since these three groups present more or less similar characteristics and distribution to the cluster results for the printed catalogue; they can be labelled according to the above titles given to the cluster results for the printed catalogue. Therefore, the digital catalogue distribution is oriented towards ‘*Black & White*’ (Cluster D1), ‘*Documentation*’ (Cluster D2) and ‘*Practical & Illustration*’ (Cluster D3).

Table 5 provides a snapshot of the five components and three clusters for the printed and digital catalogues. The numbers of pivotal editorial houses characterising the five possible strategic catalogue orientations are reduced in the shift from the printed to the digital catalogue. As a result, the association between the results of the PCAs and the cluster analyses bring to light a correspondence but also a tendency towards more specialisation regarding digitisation strategies. The weight of didactic and educational publishers (Components 4 and 5 for the digital catalogue) are reflected in Cluster D1 only. It is interesting to stress that the willingness to specialise is corroborated by the discovery that the digital PCA did not disclose multiple strategic moves endorsed by the same editorial houses, as was the case for the printed PCA (see Pearson ERPI, Eyrolles and Nouv éd de l’université).

Table 5: Schematic representation of Components 1-5 and Clusters 1-3, for printed and digital catalogues

<i>Printed catalogue</i>				
Component 1 <i>‘General sciences, literature and technical documentation’</i>	Component 2 <i>‘Youth and comics’</i>	Component 3 <i>‘Practical documentation’</i>	Component 4 <i>‘Computer science’</i>	Component 5 <i>‘School publishing’</i>
Flammarion Seuil Pearson ERPI PUF	Casterman Dargaud Dupuis	Nouv. éd. de l’université Eyrolles	Eyrolles Hermès science publications Pearson ERPI	Nouv. éd. de l’université Pearson ERPI

¹³ Dupuis, Le Lombard, Bamboo, Dargaud, Guides de Voyage Ulysse and Nouvelles éditions de l’université.

Cluster P1 (4 eds) 'Black & white' Flammarion Seuil Pearson ERPI PUF	Cluster P2 (20 eds) 'Practical & Illustration' Casterman Dargaud Dupuis Nouv. ed. de l'université Eyrolles Hermès science publications	Clusters P1 ('Black & white') and P2 'Practical & Illustration'		
Cluster P3 (76 eds) 'Documentation'				
CNRS editions Presses de Sciences Po Presses polytechniques et universitaires romandes Lavoisier Payot Rivages				
Digital catalogue				
Component 1 'General and social sciences, informatics and technical documentation'	Component 2 'Youth, arts and comics'	Component 3 'Practical documentation'	Component 4 'School & language editions'	Component 5 'General literature'
Eyrolles	Dupuis Dargaud Le Lombard	Nouv. éd. de l'université Guides de voyage Ulysse	Pearson ERPI	Seuil
Cluster D1 (10 eds) 'Black & white' Eyrolles Pearson ERPI Seuil	Cluster D3 (6 eds) 'Practical & Illustration'		Cluster D1 ('Black & white')	
	Dupuis Dargaud Le Lombard Nouv. éd. De l'université Guides de voyage Ulysse Bamboo			
Cluster D2 (84 eds) 'Documentation'				
All of the 76 eds from the Cluster 3 printed catalogue, plus others like Actes Sud J'ai Lu Fleurus Novalis				

6. Discussion and concluding remarks

At a time when the digitisation of enterprises is at the heart of the global economy and value creation, one of the particularly striking and important dynamics to be analysed concerns new forms of differentiation that are taking place within sectors that have hitherto shared the same characteristics. The differentiated manner in which companies are adapting to the new environment created by digital technologies by changing their positioning and development strategy is therefore particularly important to grasp. More specifically, it is important to understand to what extent the strategic choices made by these companies are simply linked to the production processes, the size and context affecting a particular industrial sector, or are also dependent on the specific contents and forms of the offers of each of these firms. Consequently, it is of great significance to take advantage of the idiosyncrasy of the CCIs, since they are firstly emblematic and precursory in the adoption and digitisation of new economic models, and secondly to the extent that competition and strategies explicitly take place on the basis of different types of content, more than on the sole basis of business models or technology. As shown by Cohendet & Simon (2015) in all sectors of the economy, companies place innovation and creativity more and more at the heart of their development strategies, drawing inspiration, for this reason, from the situations and models developed in the CCIs. From this perspective, it is well worth considering whether, in such a context of creativity, digitisation progresses in a homogeneous way and calls for comparable strategies or, conversely, leaves room for differentiated strategies according to the nature of creative content

and the activities of the firm. This question has made it meaningful to explore — in all respects — the case of books and the digitisation of catalogues by publishers.

Despite similarities in terms of value creation and the supply chain, this article highlights that the digitisation of editorial houses is not only linked to a global and monolithic strategy, but also requires specifically investigating the nature of the services and products designed in order to identify and characterise digital strategies. As a matter of fact, Koch and Windsperger (2017:5) recently underlined how digitisation goes beyond the simple use of digital technologies: the organising logic among heterogeneous firms (interconnected by a common digital infrastructure) has been affected by the pervasiveness of digital technologies, with socio-technical implications.

6.1 A strategy oriented towards specialisation

The results shed light on one of the first pieces of evidence of this splintering of hitherto monolithic strategies: the existence of different components of catalogues that contribute to shaping differentiated publisher strategies. Sticking to the PCA results, the cluster analysis for the printed and the digital catalogue distribution revealed, for each distribution channel, quite similar influences of the various categories of books: three main clusters were clearly identifiable.

These results draw attention to the interesting phenomenon of differentiation that has taken place during the transformations that have affected the book industry in the digital age. In this context, the digitisation of catalogues has been a real strategic competitive resource enabling publishers to build their own specific strategies: by seeking to consolidate the position they had in the printed market, by trying to take advantage of digitisation by relying on particularly buoyant categories of books, by developing skills and by constructing a form of learning process regarding new online markets by targeting particular book segments. Therefore, the results help to establish a deep characterisation of the strategies of editorial houses in the digital age as well as throw light on the effective influence of printed book categories on digital book dissemination.

An interesting aspect concerns, in particular, small independent distribution sale networks that are threatened by these disruptive changes in the traditional oligopolistic structure of the book market. If publishers are able to specialise in specific e-book sectors, they are likely not only to survive but also to become successful thanks to the advantages provided by digital platforms. In particular, they could reach this goal through the ability to aggregate a very large readership worldwide, on niche content and markets, thus escaping the constraints of printed distribution addressing buyers in a very localised geographical distribution. The results of the PCAs and the two-step cluster analyses point to this possibility as they did indeed pinpoint certain key differences among the final group distribution (i.e. printed *vs* digital), in particular in relation to the highlighted tendency towards specialisation.

From a more global perspective, three groups of publishers have been closely but distinctly identified in printed and digital catalogue distribution. While the content of the three Clusters for printed and digital catalogues is very similar, the comparison among the related groupings draws attention to a tendency towards specialisation in the shift from printed to digital distribution. It suggests that the nature of publications (and thus of the editorial process, the readership segment targeted and the physical format of books) contributes to a new structuring of the specialisation of the book market in the same way for both physical and digital media. In fact, while the Cluster groupings for printed and digital catalogues have the same labels and echo one another, they are different in terms of the number of publishers involved. More specifically, in the shift from the printed to the digital catalogue, it can be highlighted that the '*Practical & Illustration*' grouping is greatly reduced (6 for digital *vs* 20 for printed). This could be explained by a willingness towards more 'specialisation', meaning that the few publishers included in the digital cluster '*Practical & Illustration*' are focused most of all on these categories (i.e. comics and tourist guides on average) and they have clearly chosen a digitisation strategy. A similar reasoning can be made for the '*Black & White*' category, but this time the digital cluster is larger than the printed one (10 for digital *vs* 4 for printed) meaning that this book sector is globally strongly digitised. Finally, a specific tendency towards '*Documentation*' digitisation is also observable: Cluster D2 (digital catalogue) is the mirror of Cluster P3 (printed catalogue) but it is even larger because of the acquisition of some publishers previously included in Cluster P2 (84 for digital *vs* 76 for printed). This finding is confirmed by the PCAs: while the components identified for the printed catalogue revealed the predominance of the same pivotal editorial houses for more than one component, the PCA run on the digital catalogue identified fewer key publishers and did not duplicate editorial houses for more components.

This orientation towards specialisation reflects in some way the co-specialisation strategy of complementary products in the digital market, like video games and video game consoles (Benghozi, Chantepie, 2017; Grant, 2016; Cennamo, Santalo, 2013). Applied to the book publishing market, this means that the specialisation

strategy of editorial houses in digitising their catalogue could lead to a differentiated distribution of different types of books, leading to a situation where some categories of books could be more adapted for reading on specific e-readers and some e-readers could be designed to accommodate the reading of specific categories of books and not others. More precisely, the stability of strategies in both print and digital catalogues also suggests that different clusters should experience and manifest different rates of evolution towards digital products, more or less at high speed. Verification would involve completing the analysis of catalogue data with that of commercial sales. However, as we underlined in the introduction of the article, this was not the choice made at the outset, for analytical reasons and because of the available data.

In fact, in confirming the migration (and consequent specialisation) of book categories such as practical and consultation books towards the e-book format, our analysis substantiates that the digitisation efforts of publishers are not carried out in a unique and undifferentiated way throughout the book industry. The strength of such efforts is dependent, to a considerable extent, on the content and nature of the books.

6.2 *Dynamic complementarities*

The different choices of catalogue digitisation underscore some essential concerns about the adoption of strategies in relation to technologies, the principal strategic means through which a system can obtain an early and decisive advantage over the competition (Mangematin, Callon, 1995: 456). The book publishing industry is in a privileged position: the race for innovation that leads companies to outdo each other in order to ultimately retain their initial position is carried out in an oligopolistic and less competitive environment. The publishing industry was not 'born digital' like the video games sector (Benghozi et al., 2017): this aspect could make a difference and partly explain this rigidity and 'path dependence' (David, 1985) on its age-old traditional structure, the tendency to follow and to adapt to the 'fashion of the moment' and the difficulties in renewing its strategy in line with the new digital environment.

To this extent, the strategy followed by our sample of publishers can be understood in terms of dynamic complementarities. According to Love et al. (2014: 1774), '*two discrete activities are complementary if adding one activity increases the returns from doing the other. This implies that the benefit of adding a new activity depends not simply on what the firm currently does, but on what it did in the past: it concerns adding something to an existing strategy*'. The cluster analysis results for the printed and the digital catalogue highlighted the same number of clusters that reflect one another and that differ only in the number of publishers inside each single cluster. This result could highlight an attempt to maintain the same strategy for the printed catalogue and for the digital catalogue, but with some minor changes that could be interpreted as a sort of strategy of dynamic complementarities. In other words, if a publisher makes the choice of digitising its printed catalogue, a completely different and original strategy for the digital catalogue is not identifiable. This is further confirmed by the assumptions of Riot (2013) about the comic book category. Riot (2013) focused on the French book publishing sector in the era of e-book disruption when she analysed the evolution of the perimeter of activities of French editorial houses: the expansion of the book sector in categories like comics and practical guides was underlined. Comics in particular are among the book sectors that seem to benefit from a participative logic: they take advantage of support from a large community of fans who share free and payable models. Comics represent a smaller segment than the literature category in the French book market, but are still considered a very dynamic book sector.

The different features of the book industry at the country level, due to national identities or habits and reflected in the book catalogues, overcome the commonalities provided by international technological standards. In other words, national industries are coping in different ways with the same waves of technologies. The current situation, in the now-networked digital economy, is characterised by the fact that technological dynamics are potential levers of convergence or divergence in the market content. Given the weight of the demand side (consumers' choices) compared to the supply side (publishers' influences on consumers), we could expect that the globalised technical standards will simply complement this framework with an impact on investment strategies. The pace of technological change requires massive investments to deploy the necessary infrastructure and the appropriate content production.

This situation explains the paradoxical alternative reflected in publishers' strategies in the face of the digital age. On the one hand, the book publishing industry is witnessing a sort of continuation of the 'casino strategy' at play in most of the CCIs, characterised by a high-risk level and uncertainty of success and, correspondingly, by growing marketing and distribution expenses. On the other hand, however, in order to reduce risk and ensure successful revenues, publishers also strive to replicate successful recipes or follow current trends. The editorial houses' catalogue digitisation strategy is driven more by the easy imitation of one another than by a voluntarist drive to influence consumers' needs.

This desire to reduce risks is all the stronger as digitisation has led to a significant increase in investments by publishers to adapt to new distribution and reading channels (e-readers), to convert their catalogues to different digital formats and to recruit new staff with IT skills. Consequently, we observe a mix of radical and incremental innovations carried out by different but closely related actors, namely e-reader producers and editorial houses, all working in the same industry. The editorial houses' strategy is to respond slowly to the digital revolution in catalogue digitisation with an approach guided more by a pragmatic adaptation to the needs of the moment than by a singular long-term vision. The very limited nature of the development of the e-book in France, as well as the recent failure of a collective common platform of French booksellers¹⁴ (Poirel, 2015), indicate that this slow and conservative strategy might be successful. From this perspective, the slight tendency towards specialisation highlighted by the PCAs and cluster analyses groupings (i.e. printed vs digital) suggests that 'differentiation' could be the key to success.

6.3 Limitations and future research

In conclusion, we can therefore affirm that the empirical results obtained in the French case from a very broad base of publishers fully confirm our initial hypotheses, which were based on the assumption of a progressive differentiation of strategies in the book industry under the effects of digitisation. Additionally, we have shed more light on the factors on which this differentiation is based, such as specialisation in certain book segments, the complementary nature of printed and digital catalogues and risk minimisation.

This empirical investigation is naturally not without some limitations. First of all, the population of editorial houses is somewhat limited in number. Secondly, the study is restricted to the French market: the *Electre* databank did not provide the opportunity of an international cross analysis with other European databanks. Thirdly, the analysis was made on the basis of data collected over a specific period (2014), which is not without influence against a technological and economic backdrop that is evolving very quickly. Nonetheless, despite these limitations and potential biases, the originality and relevance of this empirical investigation are supported by the systematic quantitative analysis of publishers' catalogues and the consistency of the results with the literature and other in-depth observations. Data about e-books are not systematically available (Waldfoegel, Reimers, 2015). Given the lack of reliable official statistics in the comparison of the catalogues of printed books with those of e-books, this analysis is built on an original database extracting primary data from the most reliable and recognised French official databank in the book industry. Processing these data brought to light interesting findings, useful for better understanding the current strategies of editorial houses faced with the e-book revolution, as well as for stimulating further research.

The different dimensions of the discussion of the results open several potential avenues for future research. The first one is related to the consequence of publishers' strategies on the market for devices. It would be interesting to investigate whether the e-readers and the device market is in line with the specialisation of the digitisation of publishers' catalogues and whether this could create some specific opportunities in some market segments (namely, science and school or tourist guides). Another perspective is the role and weight of platforms in the new book value chain and market design. Currently, digital platforms play a key role in various layers, therefore interfering greatly with publishers' editorial strategies: they define the forms of monetisation and 'rules of the game' for the sector. On the one hand, the implementation and use of digital technologies make it possible to rely on the processing of digitised content and the processing of personal data, and give platforms and/or publishers the benefit of algorithms, interfaces and recommendation tools that make it possible to radically rethink prescription methods, and the books put forward on the supply side and the choices of readers and forms of access to books on the demand side. On the other hand, the mastery of these same technologies gives platforms the ability to influence the market, as already shown in the case of the music or audiovisual sectors, by establishing publishing functions and entering into exclusive agreements with certain authors in direct competition with publishers.

References

- Arnaud C., Soldo E. (2015), "Territorial Portfolio of Cultural Events (TPCE): a new managerial modality of territorial cultural events' supply", *Management International*, 19(2): 115-127.
- Barnes & Co. (2012), "Worldwide Book Publishing Industry", *Industry & Market Series, NAICS 51113*, pp. 1-108.
- Baye, M. R., De Los Santos, B., Wildenbeest, M. R. (2013). "Searching for physical and digital media: the evolution of platforms for finding books", NBER working paper series 19519.

¹⁴ 1001Libraires.com

- Benghozi, P.-J. & Benhamou, F. (2010): "The Long Tail: Myth or Reality?", *International Journal of Arts Management*, N°12-3.
- Benghozi P.-J., Salvador E. (2014) "Strategies and business models of online platforms in CCIs: convergence or differentiation in the e-book sector?", in Schramme A., Kooyman R. (Ed.), Hagoort G. (2014), "Beyond Frames. Dynamics between the creative industries, knowledge institutions and the urban context", Eburon Academic Press, Delft, ISBN 9789059728844, pp. 96-104.
- Benghozi P.-J., Salvador E. (2015) "Technological competition: a path towards commoditization or differentiation? Some evidence from a comparison of e-book readers", *Systèmes d'Information et Management (SIM)*, vol. 20, n. 3, pp. 97-135;
- Benghozi P.-J., Salvador E. (2016) "Investment strategies in the value chain of the book publishing sector: how and where the R&D somehow matter in creative industries?", *Technology Analysis & Strategic Management*, vol. 28, n. 5, pp. 568-582
- Benghozi P.-J., Salvador E., Simon J.-P. (2017), "The race for innovation in the media and content industries: Legacy players and newcomers. Lessons for policy makers from the video game and cinema industries", *Journal of Arts Management, Law and Society, Special Issue Digital Cultural Policies in Comparison*, ed. by Roberge J. and Chantepie P., 47(5): 346-360.
- Benghozi P.-J., Chantepie P. (2017), *Jeux vidéo : l'industrie culturelle du XXIème siècle ?*, Ministère de la Culture et Presses de Sciences Po, Paris. 272p.
- Benhamou F. (2004), *L'Économie de la culture*, Paris, La Découverte
- Benhamou F. (2014), *Le Livre à l'heure numérique. Papier, écrans, vers un nouveau vagabondage*, Paris, Seuil.
- Benhamou F. (2015), « Fair use and fair competition for digitized cultural goods: the case of eBooks », *Journal of Cultural Economics*, 39: 123-131.
- Cennamo C., Santalo J. (2013), "Platform competition: strategic trade-offs in platform markets", *Strategic Management Journal*, 34(11): 1331-1350.
- Cohendet P., Simon L. (2015), "Introduction to the Special Issue on Creativity in Innovation", *Technology Innovation Management Review*, 5(7): 5-13.
- Dagiral, É., Parasio S. (2010), "Presse en ligne : où en est la recherche ? ", *Réseaux*, 2-3(160-161) : 13-42.
- David P. A. (1985), "Clio and the Economics of QWERTY", *The American Economic Review*, vol. 75, n° 2, pp. 332-337.
- De Prato G., Simon J.-P. (2014), "Public policies and government interventions in the book publishing industry", *info*, vol. 16, n. 2, pp. 47-66.
- Electre.com pour les nuls (2009), *Services commercial et technique d'Electre*, First Interactive, Paris.
- Everitt B., Landau S., Leese M. (2001), *Cluster Analysis*, Edward Arnold, London.
- Federation of European Publishers (FEP), (2016), "Report of Activities, May 2015-May 2016", Brussels, pp. 1-60.
- Gilbert, R. J. (2015). "E-books: a tale of digital disruption", *Journal of Economic Perspectives*, 29(3), 165-184.
- Grant R. M. (2016), "Contemporary Strategy Analysis", 9th edition, Wiley.
- Howard N. (2009), "The Book: The Life Story of a Technology", The Johns Hopkins University Press, Baltimore, MD, ISBN 978-0-8018-9311-7.
- Hjørth-Andersen C. (2000), "A Model of the Danish Book Market", *Journal of Cultural Economics*, 24(1): 27-43.
- Hviid, M., Jacques, S., Izquierdo Sanchez, S., (2017). "From publishers to self-publishing: the disruptive effects of digitalisation on the book industry", CREATE working paper series 2017/06.
- Koch T., Windsperger J. (2017), "Seeing through the network: Competitive advantage in the digital economy", *Journal of Organization Design*, 6(6): 1-30.
- Latcovich, S., Smith, H. (2001). "Pricing, Sunk Cost, and Market Structure Online: Evidence from Book Retailing", *Oxford Review of Economic Policy*, 17(2): 217-234.
- Lebert, M. (2009), "A short history of ebooks", NEF, University of Toronto.
- Lescure P. (2013), « Mission Acte II de l'exception culturelle. Contribution aux politiques culturelles à l'ère numérique », *Culture-Acte 2*, Ministère de la culture et de la communication, Paris, Mai, pp. 719.
- Love J. H., Roper S., Vahter P. (2014), "Dynamic complementarities in innovation strategies", *Research Policy*, vol. 43, pp. 1774-1784.
- Mangematin V., Callon M. (1995), "Technological competition, strategies of the firms and the choice of the first users: the case of road guidance technologies", *Research Policy*, vol. 24, pp. 441-458.
- MarketLine (2014), "Books in France", *Industry Profile*, Reference Code: 0164-0781, April, pp. 1-32.

- Meng Q., Chen Z. (2008), "Principal Component Analysis and Cluster Analysis in Multi-regional Logistics Planning-Taking Zhejiang Province for Example", International Seminar on Business and Information Management, Wuhan, Hubei, China, vol. 02, pp. 15-18.
- OECD (2012) 'E-books: Developments and Policy Considerations', OECD Digital Economy Papers No. 208.
- Peltier S., Moreau F. (2012), "Internet and the 'Long Tail versus superstar effect' debate: evidence from the French book market", Applied Economics Letters, vol. 19, n. 8, pp. 711-715.
- Poirel C. (2015), "The analysis of a failed collective strategy. The case of 1001Libraires.com in the book business", Management International, 19(2): 64-82.
- Riot E. (2013), "L'édition française à l'heure de la rupture du livre numérique", Entreprendre&Innover, 2013/2 n. 18, pp. 31-40.
- Ritala P., Golnam A., Wegmann A. (2014), "Coopetition-based business models: the case of Amazon.com", Industrial Marketing Management, vol. 43, pp. 236-249.
- Ronning H., Slaatta T. (2012), "Regional and national structures in the publishing industry with particular reference to the Nordic situation", Journal of Media Business Studies, vol. 9, n. 1, pp. 101-122.
- Rouet F. (2007), "Le livre: mutations d'une industrie culturelle", La documentation Française, Paris.
- Rudiger Wischenbart (2014), "Global eBook. A report on market trends and developments", Update Spring 2014, Rudiger Wischenbart Content&Consulting.
- Salvador E., Simon J.-P., Benghozi P.-J. (2019), "Facing disruption: the cinema value chain in the digital age", International Journal of Arts Management, vol. 22, n. 1, pp. 25-40.
- Simon J. P. (2014), "Ce que le numérique a déjà changé à l'édition", INAGlobal, 21 November 2014, <http://www.inaglobal.fr/edition/article/ce-que-le-numerique-deja-change-ledition-7977>
- Simon J. P. (2015), "Amazon déclare la guerre des contenus", INAGlobal, 11 February 2015, <http://www.inaglobal.fr/numerique/article/amazon-declare-la-guerre-des-contenus-8103>
- Simon J.-P., Benghozi P.-J., Salvador E. (2015), "The new middlemen of the digital age: the case of cinema", info, vol. 17, n. 6, pp. 97-115.
- SNE (2016), "Repères Statistiques. France et International, 2015-2016 », SNE (Syndicat National de l'Édition), June, pp. 1-119, Paris.
- Tabachnick B. G., Fidell L. S. (2013), "Using Multivariate Statistics: Pearson New International Edition", Pearson Education.
- Waldfoegel, J., Reimers, I. (2015). "Storming the Gatekeepers: Digital Disintermediation in the Market for Books", Information Economics and Policy, vol. 31, issue C, pp. 47-58.

Annex 1: Electre databank¹⁵

Electre is the trading company of the Cercle de la librairie, an institution active since 1847 alongside book professionals. It offers particularly rich bibliographical information while observing the standards governing cataloging.

The Electre databank was created in 1985 and includes (in 2014, when our data were collected) 1,200,000 titles. The information for each title is provided: author(s), title, availability, price, codes ISBN, EAN, ISSN, editorial house, distributors, collections, date of release on the market, description (weight, size), abstract (written by the catalogers of Electre), abstract of the editorial house, back cover biography of authors, critics, glossary, theme, genre, Dewey levels, RAMEAU (language indexing for subject), table of contents.

Annex 2: The Dewey decimal classification¹⁶

The Dewey library classification system was conceived in 1873 and first published in the USA in 1876. It is a decimal numerical encyclopedic classification system used for marking and classifying books in a library. It is the most widely used classification system in the world.

The Dewey classification has been revised and expanded through 23 major editions: the latest edition was issued in 2011. It is currently developed, maintained and regularly updated by the Online Computer Library Center (OCLC), a library research center.

¹⁵ <http://www.electre.com/ElectreHome.aspx>

¹⁶ <http://www.oclc.org/dewey.en.html> <http://www.oclc.org/en-CA/dewey/features.html#print>

The classification's notation of the Dewey system uses three-digit Arabic whole numbers for the main classes, with fractional decimals allowing expansion into sub-classes and divisions for further details.

At the more general level, the Dewey hierarchical classification is divided into ten main classes. Each main class is further divided into ten divisions. Each division is then divided into ten sections. The ten main classes are the following:

- 000 Computer science, information & general works
- 100 Philosophy & psychology
- 200 Religion
- 300 Social sciences
- 400 Language
- 500 Pure Science
- 600 Technology
- 700 Arts & recreation
- 800 Literature
- 900 History & geography