Reading indicators on the social networks Goodreads and LibraryThing and their impact on Amazon

Nieves González-Fernández-Villavicencio (Sevilla)

Summary: The aim of this paper is to identify relations between the most reviews and ratings books in Goodreads and LibraryThing, two of the most impacting social networks of reading, and the list of top-selling titles in Amazon, the giant of the distribution. After a description of both networks and study of their web impact, we have conducted an analysis of correlations in order to see the level of dependency between statistical data they offer and the list of top-selling in Amazon. Only some slight evidences have been found. However there appears to be a strong or moderate correlation between the rest of the data, according to that we propose a battery of indicators to measure the book impact on reading.

Keywords: Goodreads, LibraryThing, reading social networks, virtual reading clubs, reading indicators, Amazon

1 Epitexts, social reading networks and promotion of reading[1]

It is acknowledged that social media in general has become a way of communicating to share a whole series of habits, behaviours and tastes, including reading and sharing books. This is the context in which Lluch et al. (2015: 798) deploy the concept of *epitext* and Jenkins's notion of interactive audiences, which are fostered by the social web and refer to groups of readers whose attention is focused on books and reading-related issues. «What we have are virtual identities for which it is equally important to keep abreast of the latest publishing releases and to exchange knowledge and opinions about books that they read, authors whom they like, themes and so forth» (Lluch et al., 2015: 798).

1 Research financed by «Virtual spaces for the promotion of books and reading. Formulation of indicators to evaluate its quality and effectiveness», FFI2015-69977-R (MINECO/FEDER), R & D & I Projects of the National Programme for Research Aimed at the Challenges of Society. Ministry of Economy, Industry and Competitiveness. Government of Spain, 2015.

ISSN 0932-2221 · eISSN 2199-7276
https://doi.org/10.46586/ZfK.2019.143-167

Social media platforms of this kind have great power to promote books and reading, as they allow the inclusion of comments, opinions, favourite books and reviews. «Children and young people are increasingly demanding to share the works that they are reading and to find out what their peers think about them. In addition, social networks can help to create expectation or interest in specific works, and editors can see the views of readers first hand and then adapt their catalogues or even customize them» (García Rodríguez, 2013: 15). Accordingly, interactive mechanisms involving the Internet, social media and the voices of *booktubers*, bloggers and writers can be used to promote books.

Many platforms have been created specifically for readers to share their tastes and impressions. In this area, Goodreads is the essential website. An article by Lluch et al. (2015: 799) states that the most widely used and well-known epitexts are blogs, commercial websites, reading forums, social networks, wikis, *booktubers* and trailers for books. However, within social media platforms, there are spaces specially dedicated to reading-focused user interaction – for example, noninstitutional virtual reading clubs such as Goodreads and LibraryThing, where creation of and access to information and social reading are organized. This is a relatively new phenomenon that has been around for no more than 10 years: Goodreads was created in 2007, Shelfari in 2006, LibraryThing in 2005 and Babelio in 2007. In Spain in recent years, various projects have emerged, such as Lecturalia in 2006 and QuèLlegeixes? in 2008 (Llobet Domènech et al., 2016: 3).

Virtual social networks and bookmarking sites like *Delicious* currently focus on recommendations and the creation and sharing of notes in relation to book lists. The results of Kaplan's study (2016: 1) suggest that this category of social reading does not appear to be very different from other forms of social media interaction. These virtual communities for social reading share not only ideas and thoughts about books and reading, but also feelings and emotional reactions surrounding these ideas, through which they acquire social networking functions.[2]

This huge amount of information that comes directly from readers ought to be very useful for companies in the sector, since it would allow them to put forward new projects and innovations. The information on reading preferences provided by these networks has great value in relation

2 As an example, Grupo Planeta has launched the project Oh!Libro (<www.ohlibro.com>), a platform where users can find their next book through feeling-based assessments made by other readers.

to marketing issues within the book industry. Users on both networks are a set of consumers of books and potential buyers, and publishers and distributors should be interested in exploring the data from these networks – what users read and what attracts them – since this information can be found in their recommendations (Laspa, 2013: 146).

Amazon is behind both platforms. It has owned Goodreads since March 2013, and it has held a minority share LibraryThing since 2006. Following Amazon's acquisition of Goodreads, with its 50 million reviews, the commercial giant consolidated its leadership in the book sector, and it has a source of first-hand information that it can use to learn about its users' tastes (Rivera, 2014: 1). On LibraryThing and Goodreads, the community receives and provides ratings and reflections and participates in debates, all of which allows profiles to be identified and reader models to be constructed. As a result, Amazon has access to privileged forums from which it can observe the entire reading process that may end up in the purchase of the book (Llobet Domènech et al., 2016: 13; Albrechtslunf, 2017: 2). Amazon's acquisition of Goodreads in 2013 did not come about without controversy and rejection by the community of users because of a loss of freedom and because their contributions would be subject to exploitation by the company (Albrechtslunf, 2017: 2).

In this study, we consider the possible impact of these reading networks for the book industry. As reading social networks with a verifiable existence as book clubs and virtual communities of readers, and behind which Amazon lies, there are reasons to anticipate a certain degree of dependence in relation to user activity on these networks and Amazon.

Our study focuses on the one hand on these networks and on the other on the relationship that may exist between readers' rating and reviewing activities on these specific social networks and the best-selling books on Amazon.

Studies that seek to determine the link between books purchased and books read or cited are a constant presence in the literature. Cabezas-Clavijo et al. (2013: 1237) have investigated whether the most borrowed books in libraries were also the most cited. In their study, they highlight that there is no correlation between the books most frequently borrowed from the University of Granada's library and the most cited books on either WoS (Web of Science) or Google Scholar.

In our case, we wish to ascertain whether the best-rated books on these networks are bestsellers on online distribution giant Amazon or even the most borrowed titles in Spanish public libraries. The object of our investi-

gation is to discover whether there is any truth to such an assumption. It is true that what the bestsellers are might vary if we looked at the books in other stores, or on other Amazon sales lists from different periods, but there is no doubt that, by volume of sales and through its connection with the two virtual social reading networks, Amazon is a good reference point.

Our research questions are as follows:

- Do the users of the two networks behave in similar ways?
- Do the most reviewed and/or best-reviewed or best-rated books on these networks correspond to the bestsellers on Amazon?, and Ccan significant correlations be established?
- Which indicators from these networks best reflect a book's impact?

2 Description and impact of Goodreads and LibraryThing

We begin with a descriptive section that seeks to specify the properties, characteristics and important features of the two social reading networks, namely Goodreads (hereafter GR) and LibraryThing (hereafter LT). We will look at the characteristics of the two networks based on the literature and on a study of their websites, and we will analyse their impact through web indicators and their usefulness for the purposes of *altmetrics*.

2.1 Description of the two networks

We based our description of the networks (Table 1) on the information provided on GR by Manso (2015), Desrochers (2016) and Thelwall / Kousha (2017), and that provided on LT by Richards (2013) and Laspa (2013). We used information provided by Domènech (2016) in relation to both networks. This information was complemented by data obtained from the sites themselves.[3]

	Goodreads.com	LibraryThing.com
Description	Goodreads.com is the biggest social network in the digital community for discussing reading. According to Thelwall / Kousha (2017: 972), it is a hybrid of a	LibraryThing.com is a social cataloguing and communication web tool focused on books. It is based on information stored in a personal library.

3 Goodreads web service (<https://www.goodreads.com/about/us>) and LibraryThing web service (<http://www.librarything.com/zeitgeist>).

	traditional website focused on books and social network activities.	
Modality	Asynchronous	
Creation year	2006 United States	2005 United States
Audience	General public (over 13 years of age). Network oriented to American readers. The site is in English, but there are groups in other languages such as Spanish. There is virtually nothing in Catalan.	English-language community, but versions in other languages are offered, such as LibraryThing.es.
Catalogue	1.5 billion books 50 million user reviews	114,806,203 books
Users	55 million users	2,174,704 users
Thematic scope	Literature in general	
Organizers	Platform created by Otis Chandler and Elizabeth Khuri Chandler, headquartered in San Francisco (California). In 2013, it was bought by Amazon and linked to its Kindle ebook service.	Developed by Tim Spalding in order to categorize his own books. *Abebooks* (owned by Amazon) acquired a minority share in 2006, and in 2008 so did *Bowker* (owned by the *Cambridge Information Group*).
Technologies used	Blogs, social networks	
Features or services that they offer	Rating and commenting on books, connecting with other readers and authors, receiving reading recommendations based on ratings, adding books. Allows bibliographic records to be imported and exported. Users can select books from Goodreads' catalogues and organize them on their own shelves and reading lists. Records can be imported manually or with an ISBN number, but the system for importing books that are not in its catalogue is complicated and clumsy. Offers recommendations for specific book pages that are generated automatically or by readers who also liked them.	Shared cataloguing of books to incorporate them into a personal library with comments and ratings on their interest to readers. The books are both ones that users own or have read and ones that they want to acquire or read. Includes search capabilities and incorporates bibliographic information from Amazon and 1,051 libraries using the Z39.50 protocol, which allows users to access bibliographic information through the Dublin Core Protocol and MARC format. The book import feature is very powerful. Users can select and import books from different databases from around the world, as well as musical recordings and films.
Participation tools	Recommends books based on the customized information about a reader on the network (reviews, ratings, authors, etc.), and sometimes links to videos relating to the books being looked at appear.	After incorporating the basic description of the book, users can review, rate and tag according to their own criteria. Users can access author profiles from any work and fill in information.

	Offers discussion groups for commenting on books (book clubs, magazines, student groups, etc.) and the possibility of opening other groups. Also features games related to the book world (competitions, bibliography-based activities, etc.) and provides reports on new publications that are supplemented with rankings of novels, interviews with authors and even literary awards.	Based on readers' own catalogues, it suggests readings and recommendations while taking into account the libraries of other users with similar tastes. Users can also see other users who have a similar library and leave comments as a way to contact people with similar tastes. Offers chat groups (*LibraryThing Groups*) to exchange views.
Unique aspects	Publishes monthly bulletins of new releases and has a program for authors to promote their works. Offers a space called *Listopia* [4] that allows users to make lists to indicate the works that they have read, give them a rating and indicate future readings. These lists are dynamic, and the order of the works on them changes based on user votes. Produces an annual selection of the year's best books that is based on reader recommendations and called *Goodreads Choice Awards*. The award has achieved high participation levels. Through the *Goodreads Choice* literary awards, the best books of the year are chosen by user votes, with millions of votes often being cast. Regularly organizes real-world meetups to swap books or for literary pub crawls. Provides links for purchasing books on Amazon. Links to GR reviews can be found in many library catalogues.	Features review competitions that let users acquire books or appear on a list of the 25 most prolific reviewers or of the community's 50 best-rated authors. All versions of LT offer the possibility of accessing a full page of statistics on the current scene, with listings and detailed information.[5] Its team includes librarians, and there is a specific section for this professional sector called *LibraryThing for Libraries*. Has an app called *Readar*, which shows bookshops, libraries and book-related events near users' geolocations. Members can contribute information about events. Accounts are free for up to 200 references. From that point, there is an annual or lifetime fee, though payment of the fee does not mean that users do not have to view advertisements. It is true that the design is rather outdated and unclear, and the amount of information provided is excessive.

Table 1. Goodreads and LibraryThing features.

4 Listopia web page: <http://www.goodreads.com/list>.
5 LibraryThing statistics page: <http://www.librarything.es/zeitgeist>.

In an article by Asadi et al. (2017: 104), GR and LT occupy second and third positions in a ranking that compares up to twelve social reading networks according to 9 criteria evaluating technical aspects, usability and security.

In 2017, Amazon introduced a new weekly list, Amazon *Charts*,[6] in which books are classified by the most purchased (or best-selling) and the most read at that particular moment. Books sales here encompass printed, digital and audio books, but reading is based only on e-books and measured through *Kindle*. Reading data for print editions may well have been collected from GR, but there has still been no specific analysis of the site based on data, despite the existence of publications on GR and the commercial success and number of book ratings (Thelwall / Kousha, 2017: 982). On the other hand, although GR is a site with a large audience and high commercial value for publishers, that value and its open nature expose it to spam in the form of fake reviews, and this presents a problem for Amazon.

2.2 Web impact

We will now address the issue of measuring the impact of these platforms through web indicators. We will also look at the role played by these platforms in relation to *altmetrics* and their use as recommendation tools for libraries. Finally, we will describe the indicators that the two platforms provide.

2.2.1 Indicators for measuring the platforms' web impact

To measure these platforms' impact on the web, generally search engines that count occurrences and mentions of the object are used, so that those with more impact receive more mentions.

Google Trends: interest in the two platforms over time is very different. LT is more limited and niche, whereas interest in GR has scarcely fallen.[7]

6 Amazon Charts website: <https://www.amazon.com/charts>.
7 GR and LT on Google Trends: <https://trends.google.es/trends/>.

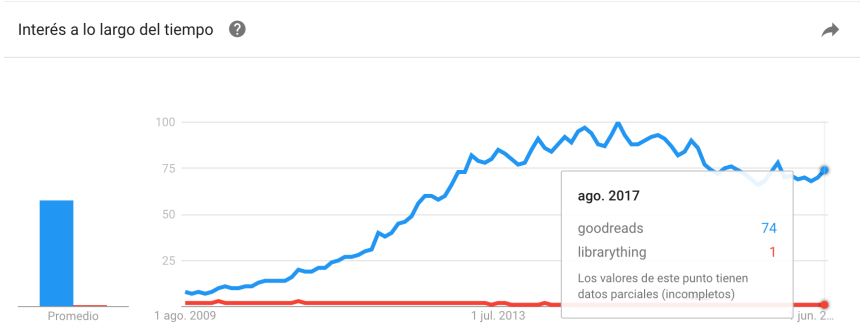

Figure 1. Evolution of Goodreads and LibraryThing in Google Trends.

To conduct a web analysis of these networks, we examined data provided by certain tools that allow estimates of the validity of the applications to be produced, based on the number of links to them and the quality of those links according to the importance of the pages that they come from (García-Carretero et al., 2016: 499). We will describe the tools, and the data are provided in Table 2.

Alexa estimates the popularity of a site using its global rank.[8] *Moz*[9] is a company that focuses on developing software for analysing SEO indicators. We used the free-of-charge version of *Open Site Explorer*, which provides information through two main indicators: *Domain authority* and *Page authority*. The *Domain authority* indicator provides a summary of a site's importance and refers to the entire domain. It consists of a 0 to 100 scale, and it can help to predict the position of a website within search engines, allowing it to be compared with others.

Moz calculates this indicator (but does not provide its formula) through a combination of different link metrics in a single score. *Page authority* measures the ranking strength of an individual page, and its score helps to predict the location of a particular page in search engines, based on link analysis carried out by the company. The combination of DA and PA is pertinent because it makes it possible to understand the potential for visibility (or authority) of the entire domain and not only of a page.[10]

8 Alexa's global rank for GR (<http://www.alexa.com/siteinfo/goodreads.com>) and LT (<http://www.alexa.com/siteinfo/librarything.com>).

9 Moz website: <https://moz.com>.

10 Comparison of GR and LT on Moz: <https://moz.com/researchtools/ose/comparisons?site=https%3A%2F%2Fwww.goodreads.com%2F>.

Majestic[11] has developed its own database for link analysis. We used its free-of-charge application *Site Explorer* and the indicators *Trust Flow* and *Citation Flow*. *Trust flow* indicates the importance or quality of a site. It estimates the distance from a domain to a series of reference domains according to quality and proven reputation. Sites that are closely linked to a site with reliable origins receive higher ratings. At the other end of the scale, sites that are far from reliable sites receive lower ratings. *Citation flow* is used to measure the number of links or the powerfulness of the website in question according to the number of sites linking to it.

The data provided by the different tools for the two social reading networks appear in Table 2. As can be seen, there is a certain amount of consistency between them in terms of the number and quality of links to a page or site. For all measurements, the values are higher for GR than they are for LT.

Indicator (0 to 100 scale)		Goodreads.com	LibraryThing.com
Alexa Rank. Site popularity		364	22,934
Moz	*Domain authority*: indicates a site's importance.	94	83
	Page authority: indicates the page's authority in search engines.	93	86
Majestic	*Trust flow*: indicates the importance or quality of a site.	81	52
	Citation flow.	67	55

Table 2. Table summarizing the web indicators for the two platforms (data obtained on 14 June 2017).

2.2.2 Impact on social media, Goodreads and *altmetrics*

In the field of *altmetrics*, nontraditional metrics for scientific output, the data offered by both GR and LT are highly relevant when measuring the impact of books (Erdt et al., 2016: 1147). GR seems to be a reasonable

11 Majestic website: <https://es.majestic.com>.

source for this objective, since it includes a large number of reviews and ratings by users and readers from inside and outside of academia (Kousha et al., 2017: 2005). Its set of indicators is a good reflection of the impact of a book in relation to other metrics, especially in the fields of human and social sciences, and by content type (Zuccala et al., 2015: 334). Specifically, the number of reader ratings on GR can be used as evidence of the value of an academic book, as has been demonstrated by some studies of correlations with citation systems (Kousha / Thelwall, 2015: 728).

Thelwall / Kousha's (2017: 974) study shows weak but statistically significant positive correlations between GR ratings and Scopus citations. However, it is important to take into account that fiction is more frequently found on GR than academic books are.

Through its *PlumX* product, one of the altmetric data aggregators, *Plum Analytics* (EBSCO) includes GR ratings, reviews and readers.[12] *PlumX* has recently been incorporated into the article metrics offered by SCOPUS. Torres Salinas et al.'s (2017: 9) article demonstrates that *PlumX* is a good indicator of book impact, but there is also a moderate correlation between the number of reviews on Amazon and on GR, which was previously confirmed in other studies (Kousha et al., 2017: 2005; Thelwall / Kousha, 2017: 974).

Kousha et al. (2017: 2005) have worked with a GR indicator called *Engagement Count*, which is the sum of the number of reviews of a book; the ratings; the times that it has been included in a list of read books, books to be read or books being read; or the times that it has been added to a collection. This indicator has a low or moderate correlation with the rest of the scientific and non-scientific impact indicators, with better results in the areas of arts and humanities and some social sciences, and it broadly suggests that the feedback of readers on GR reflects multiple types of intellectual impact, including that of a scientific, educational or cultural nature, as this kind of *altmetrics* often reflects.

2.2.3 LibraryThing's impact on libraries

There are two reasons why libraries use LT. First, they seek to promote their collections, in the case of university libraries in relation to new acquisitions, and to a lesser degree in the case of public libraries. Second, it

12 GR indicators included by PlumX: <http://plumanalytics.com/plumx-adds-further-book-support-with-goodreads-metrics/>.

allows them to engage users of the library and of LibraryThing, especially in the case of school libraries, through book recommendations and the addition of comments and tags (Richards / Sen, 2013: 499). Broadly speaking, libraries look to increase access to and dissemination of collections through their use of LT. However, there are no general data about the use of these accounts by libraries to ascertain if users are accessing these profiles, what use they make of them, and whether all this implies an increase in loans from library collections.

2.2.4 The platforms' own indicators

These platforms provide their own indicators and statistics. These are described in detail by Montesi / Esteban Aragoneses (2014: 226) and will be the object of our study (Fig. 2 and 3).

GR, LT and Amazon all display users' average rating as a main indicator. For each title, GR (Fig. 2) shows a set of statistics indicating how many times and when a title is added to a library (*add by people*), the number of *ratings* and *reviews*, and whether it has been marked as read or to be read. These data refer to the six months prior to the search.

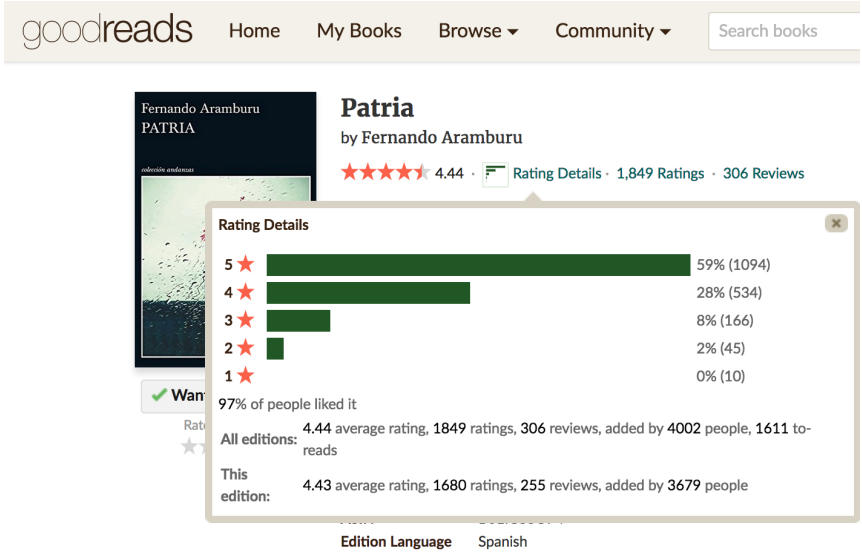

Figure 2. Goodreads statistics for *Patria*.

LT's statistics (Fig. 3) are less structured than the previous ones, except for the indicator of *popularity*. This indicator shows the position of a title in a ranking, taking into account the number of copies catalogued between its first appearance and the present. The indicator *mentions* compiles all discussions in which the book is mentioned. The data are updated every 24 hours. On these networks, the *Reviews* indicator appears regardless of the language. *People* indicates the number of users who have added a book to their personal library.

Figure 3. LibraryThing statistics for *Patria*.

A book's impact on GR is usually reflected by the number of additions to a personal library and the number of reviews published. According to Montesi / Esteban Aragoneses (2014: 237), the former may be a reliable reading indicator since the personal collection determines the contacts that the system suggests and recommendations for future readings, although it does not necessarily mean that the book has been read. On the other hand, the number of reviews is probably the indicator that reflects reading activity with the greatest confidence since, according to these authors, a review is not published openly if the book has not previously been read. For Halevi et al. (2016: 199), the most important indicator is book mentions. Kousha / Thelwall (2015: 728), meanwhile, recommend the number of reader ratings on GR as evidence of the impact of an academic book.

In the case of LT, the greater the number of users who have added a book to their personal library (*People*), the better the social aspects of the ratings, recommendations or tags will be (Laspa, 2013: 152).

On Amazon.es, for each book a classification is provided according to the position that it occupies within different thematic lists of best-selling books, either from the previous year, top sellers at any time, or those that are selling best at the time of the search. Through this dynamic lists that

are updated every hour are created. In our study, we have preferred this latter option, since what we were interested in was to see how impact on social reading networks was reflected in immediate sales. We preferred print versions to *Kindle* versions, provided that the print version occupied a higher position on the lists. When a same book appeared in both versions, we only chose the print version.

3 Methodology

Based on the descriptive analysis performed on both reading networks, we used the dynamic lists of best-selling books on Amazon.es at that moment in time in the area of literature and fiction, as well as books' average ratings given by users, to ascertain the impact of a set of statistical indicators: for GR, *People*, *Reviews*, *Ratings* and *Average Ratings*; for LT, *People*, *Reviews*, *Popularity, Average Ratings* and *Mentions*. We worked on the basis of direct observation and then implemented statistical analysis techniques using Excel and SPSS to see if there was any correlation between these data.

We wish to point out that for the purposes of data collection in this study we used the Spanish versions of Amazon (<Amazon.es>) and LT (<LibraryThing.es>), although the latter contains the same statistical information as the *.com* version, and so we refer to the network without specifying the site. GR has no site specifically in Spanish.

3.1 Statistical analysis and correlations of variables

We captured activity indicators over two time periods in order to corroborate data. These periods were 22 August 2016 and 8 July 2017. Summer dates were chosen because this is the period when fiction, the content type that stands out on these platforms, is most widely read.

On Amazon.es, we selected two specific moments on the indicated days, since listings are updated every hour.[13] We always took the data for the edition with the highest position in the listings, which in most cases involved *Kindle* versions, the format that is most widely used for summer reading.

We obtained reader ratings of these books from GR and LT. Where a book from the list of bestsellers on Amazon did not appear on GR or LT,

13 Lists of best-selling books by categories, updated every hour: <www.amazon.es/gp/bestsellers/books/902674031/ref=zg_bs_nav_b_2_902675031>.

we eliminated it from the study, since no relations between the platforms could be established.

In terms of library lending data, the *Baratz* community has produced a compilation of the most borrowed books from Spanish public libraries in recent years. As *Baratz* is the owner of the library management system most widely used by state public libraries and many municipal libraries in Spain, the list is quite significant and brings together the most borrowed books in 2016.[14] We observed an overlap between this list and the list of best-selling books on Amazon.

3.2 Study results

In general terms, we can see a greater coincidence between the statistical results from GR and LT, and between GR and Amazon, but a much lower coincidence between Amazon and LT. A significant proportion of the best-selling books on Amazon do not appear on LT, and nor do they on the general list of best-selling works from the summer of 2016, something that does occur in the case of GR. We might ask if the behaviour of readers is different on in LT than it is on GR. In some cases, we found the original version of the work but not the translation to Spanish that appears on the Amazon list.

Most of the best-selling books on Amazon tend to be the *Kindle* version, which does not confirm the strong trend found in a September 2016 study by *Pew Research Center*[15] that asserts that printed books are still more popular than digital ones.

The statistics offered on LT correspond not only to the *.es* space, but to all of LT's spaces. Hence translated works by foreign authors have more ratings, comments and reviews than do those by Spanish authors that are only available in Spanish.

Figures 4 and 5 show a comparison of *Average Ratings* of the books studied on the different platforms, in 2016 and 2017. There is no relationship between the position occupied by a book within the bestsellers on Amazon and the *Average Ratings* by users on Amazon, GR and LT. Books

14 List of the most borrowed books from public libraries in 2016, produced by Baratz: <www.comunidadbaratz.com/blog/los-30-libros-mas-prestados-en-las-bibliotecas-publicas-de-espana-en-2016/>.

15 Report by Pew Research Center on book-reading habits: <www.pewinternet.org/2016/09/01/book-reading-2016/>.

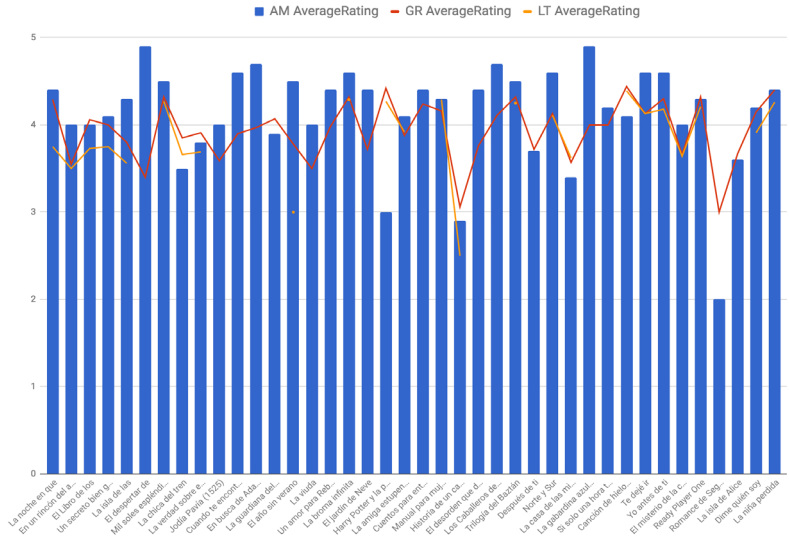

Figure 4. Average Rating Comparative Table from Amazon, Goodreads and LibraryThing (August 2016).

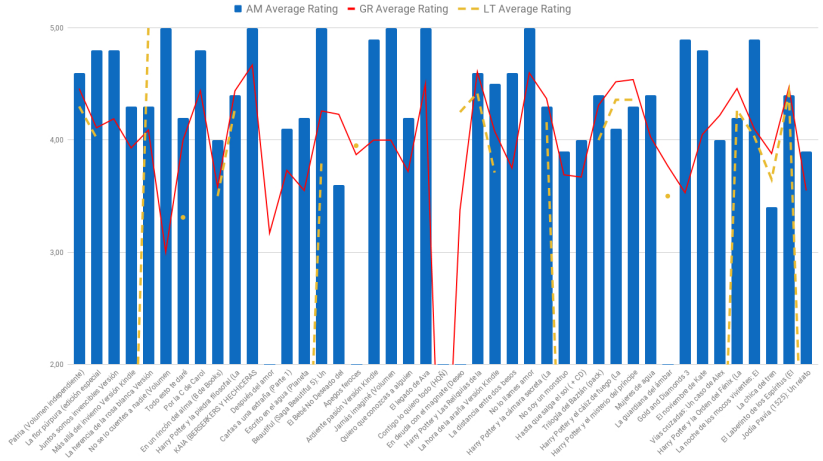

Figure 5. Average Rating Comparative Table from Amazon, Goodreads and LibraryThing (July 2017).

Figures 4 and 5 show a comparison of *Average Ratings* of the books studied on the different platforms, in 2016 and 2017. There is no relationship between the position occupied by a book within the bestsellers on Amazon and the *Average Ratings* by users on Amazon, GR and LT. Books titles appear on the horizontal axis according to the position that they occupy in the best-seller list, with number 1 appearing closest to the left. The *Average Ratings* on the three sites are distributed evenly on the horizontal line, and they do not reflect the position of the best sellers.

Nor is there a clear direct relationship between bestsellers and *Average Ratings* on Amazon. We found high values between the first and last positions and low values between the first positions. A slight tendency toward a concentration of lower values in the last positions can be observed, but only in the 2016 data, and so there is no significance. For the 2017 data, we repeatedly found that some of the best-selling books had not received any rating on Amazon, indicating that that they were recently published.

The *Average Ratings* from the three platforms moved fairly uniformly in a mean of high ratings. GR ratings were generally lower than those on Amazon, and those on LT were lower than those on GR, as in some cases titles did not even appear on that network. However, we can that that *Average Ratings* on GR and LT were fairly consistent with one another for both the 2016 and 2017 data. As we will discuss later, other factors may interfere with the high ratings on Amazon.

The data in the following figures (6 and 7) relating to GR and a on logarithmic scale show a correspondence between *Ratings, Average Ratings*,

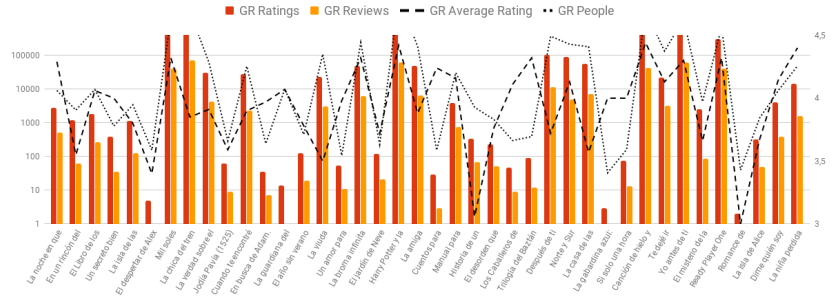

Figure 6. Goodreads Ratings, Reviews, Average Rating and People (August 2016).

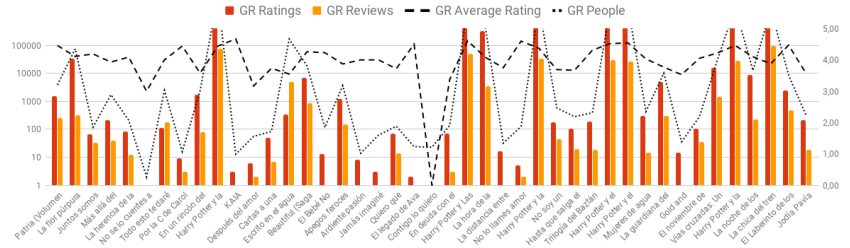

Figure 7. Goodreads Ratings, Reviews, Average Rating and People (July 2017).

Reviews and *People* who had included these books in their libraries for each of the books, with the Harry Potter books usually being the best rated and having the most reviews.

In the case of LT, and also on a logarithmic scale (Figures 8 and 9), there is coincidence for most of the books between a greater number of people who have them in their libraries and a greater number of reviews, but in general the number of reviews on LT is very low. A study of the Pearson correlations reveals moderate and strong values between the variables, which are intensified if Spearman's Rho coefficient is applied. The results of the previous charts were confirmed.

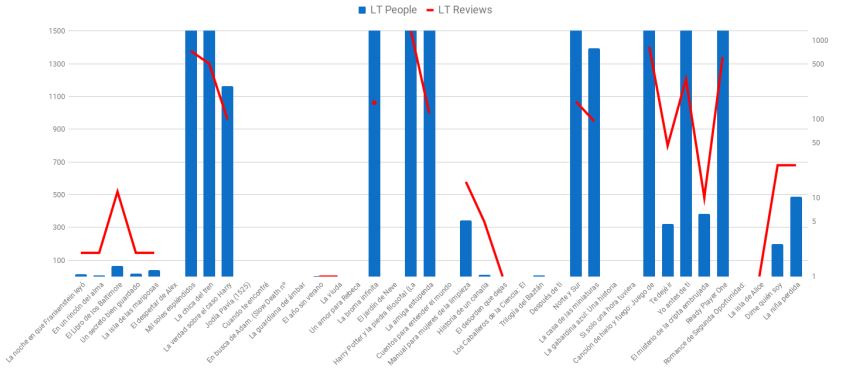

Figure 8. LibraryThing: People and Reviews (August 2016).

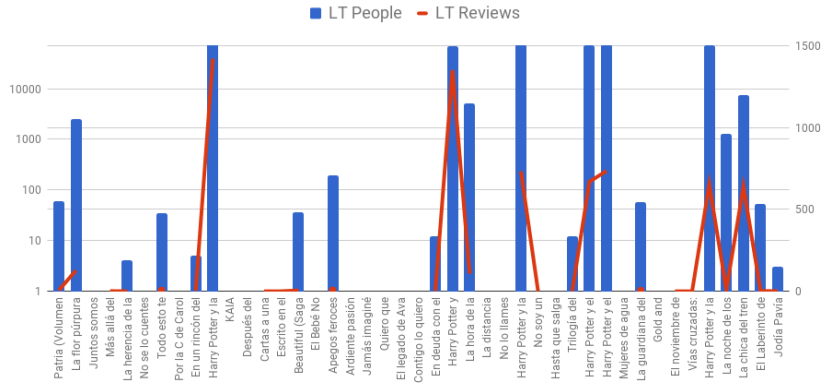

Figure 9. LibraryThing: People and Reviews (July 2017).

Amazon's ranking number, the order in the sales list, does not correlate with any of the variables from the other platforms, or even with the *Average Ratings* of users of Amazon, where the correlation is even negative, for both the 2016 and 2017 data. For 2016, it only correlates, and weakly so, with two variables from LT, namely *Mentions* and *Average Ratings*, but these values decrease considerably for 2017. Since the data from the three platforms that we have studied are cumulative, they do not exhibit a relationship with the list of best-selling books at a particular current moment of encountering titles, which may even have just appeared on the market.

The *Average Ratings* from Amazon and GR correlate moderately for 2016 and 2017. This is the only value that correlates with the *Average Ratings* from Amazon. The *Average Ratings* from GR and LT also correlate with one another, perhaps because of the close relationship between the platforms. However, these mean variables do not seem in general to have correlated with the rest of the variables. Only *Average Ratings* from GR and LT moderately correlate with *Mentions* from LT. Therefore, Amazon's *Average Ratings* only correlate with GR's *Average Ratings,* and its values are not reflected on LT. However, the *Average Ratings* from GR and LT correlate, but they scarcely do so with the rest of the variables, even within the same platform.

GR's *Average Ratings* are the ones that correlate most strongly with the other means and with some variables from LT, but they do not correlate with any from the same platform.

The GR variables of *People, Rating* and *Reviews* correlate moderately and strongly among one another, but not with GR's *Average Ratings*, as we have just indicated. They also correlate strongly with *People, Reviews* and *Mentions* from LT. They do not correlate with the variables from Amazon, except for GR's *Average Ratings*, which do so with the other *Average Ratings*.

The *People* variables from GR and LT correlate very strongly in both time periods, as do also the *Reviews* variables from the two platforms. The *People* and *Reviews* variables from GR and LT correlate moderately to strongly with one another for 2016 and 2017, as can also be seen in Figures 8 and 9.

There is no correlation between *Average Ratings* and *Ratings* on GR. The LT variables do not correlate with those from Amazon, but *People*, *Reviews* and *Mentions* do so strongly between one another and with those from GR. The values are even more significant if we look at the 2017 data (Figure 7) or if we apply Spearman's Rho coefficient. They correlate negatively with *Popularity* since it is an inverse index, and they do not correlate except in the case of *Mentions* and *Popularity* with *Average Ratings*.

LT's *Mentions* variable is the one that most strongly correlates with the rest in 2016, and it is one of the variables that most strongly correlate in 2017. It does so moderately with all the GR variables (a trend that is more accentuated in 2017) and with *Average Ratings* from LT (although this is slightly lower in 2017) and strongly with the rest of the variables from LT, except *Popularity*, with which there is no correlation.

The variable *Popularity* hardly correlates, though it does so negatively – it is a reverse index – with the rest of the variables, except with *Average Ratings* from LT, and this becomes stronger for 2017.

The data obtained from the two periods of 2016 and 2017 largely coincide and confirm these results. These data indicate that the ratings and reviews contributed to the two social reading networks correlate with one another with great frequency.

As for the list of the most borrowed books from Spanish public libraries from *Baratz*, we found that only 10% of the most borrowed books from Spanish public libraries during 2016 appear in Amazon's list of best-selling books at the time of the study and that some from that same list, almost 10%, continued to appear in the listing of Amazon bestsellers in August 2017.

The original data from this study can be found in Mendeley's data repository (González-Fernández-Villavicencio, 2017).

4 Conclusions

The rise of the Internet and social media has produced dynamic changes in reading, which in the digital world has been transformed into a communal, collaborative and participatory process, a combination of production and use (Saxena, 2018). In the field of virtual reading clubs, the comments and ratings made by other readers on these platforms[16] have become a critical element in the likelihood that a book will be purchased or borrowed from a library and read.

In this study, we have sought to ascertain whether this statement is true when one compares Amazon bestsellers with the ratings and reviews that they receive on the GR and LT social reading networks. To this end, we sought answers to the questions that we posed at the beginning of the study.

- Do the users of the two networks behave in similar ways?

At first glance, it would seem that users of both networks behave similarly, since their variables correlated moderately or strongly. However, in terms of magnitude, the number of reviews or ratings on the two networks, or the number of people who have them in their libraries, they bear no relation in reference to the same books. The values for GR for both *Reviews* and *Ratings* stand out in all the metrics, as they are much higher than those found on LT.

- Do the best-reviewed or best-rated books on these networks correspond to the bestsellers on Amazon, and can significant correlations be established?

Based on the data from the sample, the best-reviewed or best-rated books on the two networks did not correspond to the bestsellers in the two periods studied, as can be seen in Figures 4 and 5. We have only detected a moderate correlation between the *Average Ratings* on GR and those on Amazon, as well as between those on GR and LT, but not between LT and Amazon.

With regard to the lack of correlation with the list of bestsellers in Amazon, it is necessary to take into account that books that are subject to higher activity in terms of reviews and mentions on social reading network platforms are classic works of fiction rather than recent literary works

16 Information obtained from the Puro Marketing website: <www.puromarketing.com/76/27702/siguen-siendo-opiniones-consumidores-amazon-realmente-fiables.html>.

(Halevi et al., 2016: 199), which may be the books that are expected to make for summer reading and which are present in the Amazon bestseller lists.

On the other hand, it is interesting to understand how the algorithm for Amazon's list of bestsellers functions.[17] Every sale or download of a book counts toward an appearance on the list, but this occurs according to how others perform. That is, one work replaces another. This works in favour of continued and sustained sales of a book and not moments of many sales. Nor is the position affected by the number of reviews of or comments about the book on Amazon.

- Which indicators from these networks best reflect a book's impact?

Based on the study undertaken and the literature consulted, the following are the indicators on these networks that we believe best reflect the impact of a book are following: average ratings, mentions on LT, and people and reviews *Average Ratings* from GR moderately correlate with *Average Ratings* from Amazon and LT. We believe that this is a fairly effective indicator of the average score from all three platforms, taking into account that Amazon ratings are not very strict.

- Mentions on LT:

The number of times that a title has been mentioned in LT's groups and forums. This is the indicator that correlated strongly and moderately with the rest of the variables the most, and so we consider it to be an indicator with great predictive power. Halevi et al. (2016: 199) consider it to be the most important indicator.

- *People* and *Reviews* from GR and LT:

The numbers of people on GR and LT who have a title in their library strongly correlate with one another and with many of the indicators studied, and therefore we consider this to be a good reflection of the attention received by the reader. Moreover, these variables correlate moderately and strongly with the number of reviews received on the two platforms. For Montesi / Esteban Aragoneses (2014: 237), these two indicators (in reference to GR) may be reliable indicators for reading. And for Laspa (2013: 146), in the case of LT, the greater the number of users who have added a book to their personal library (People), the better the social aspects of the ratings, recommendations or tags will be (Laspa, 2013: 152).

That said, some authors have taken the effectiveness of GR's *Ratings* indicator to be evidence of the impact of academic books (Kousha / Thel-

17 <http://miguelangelalonsopulido.com/como-funcionan-lista-ventas-y-algoritmo-amazon/>.

wall, 2015: 728). Thelwall / Kousha (2017: 974) describe statistically positive correlations with Scopus citations. However, there is no correlation with GR's *Average Ratings* – as we have seen, this is a reliable indicator – because to begin with users rate the books that they include in their libraries very highly, only to then award lower ratings as they include more books (Thelwall / Kousha, 2017: 981).

The limitations of this study could be addressed by a more complete data set, or by using other lists of best-selling books from Amazon. In the list that we used, the weight in the ranking lies in recent sales, but on the social reading networks studied, scores are cumulative in nature, so the scores of classic fiction works will always be higher and those of recent literature lower (Halevi et al., 2017: 199). As we have indicated, the lists of best-selling books on Amazon work in such a way that a high place on the list does not mean a greater number of copies sold or downloaded. Rather, one position depends on others (Alonso Pulido, 2016).

On the other hand, along with other authors, we take the view that the possibilities offered by these platforms are not being exploited by users and publishers. With regard to GR, authors indicate that users are ignorant of the possibilities of GR since there is a low correlation between the number of books read and reviews submitted, and this suggests that users explore and use GR without exploiting its full potential (Thelwall / Kousha, 2017: 981).

For Montesi / Esteban Aragoneses (2014: 238), LT users do not review the books that they read and are more likely to use tags. For these authors, such people represent a type of user who does not like posting personal information, perhaps because they belong to a different generation or culture from that which actively participates on GR or Amazon. In the data obtained for our study, the number of reviews, where there are any, is much lower on LT than it is on GR. In terms of magnitude, user activity on GR or Amazon is much greater than it is on LT, as a result of which it might be thought that LT's users may engage in a different kind of activity relative to that of the users of the other two platforms. Despite these differences, their statistical data correlate very significantly.

Finally, we consider the role that statistics from GR play in relation to *altmetrics*, as a metric for book impact, to be very relevant. To be sure, as these metrics become part of researchers' workflow, these social reading networks will become more important to the book industry.

References

Albrechtslunf, Anne-Mette Bech (2017): «Negotiating Ownership and Agency in Social Media: Community Reactions to Amazon's Acquisition of Goodreads», *First Monday* 22:5, 1–11.

Alonso Pulido, Miguel Ángel (2016): «Cómo funcionan la lista de ventas y el algoritmo de Amazon», in: *miguelangelalonsopulido.com* 8:4, <http://miguelangelalonsopulido.com/como-funcionan-lista-ventas-y-algoritmo-amazon/>.

Asadi, Saeid / Nourmohammadi, Hamzehali / Ardali, Massoomeh Omidi (2017): «Evaluation and Ranking of Book Social Network Websites», *International Journal of Information Science and Management* 15, 95–108.

Cabezas-Clavijo, Álvaro / Robinson-García, Nicolás / Torres-Salinas, Daniel (2013): «Most borrowed is most cited? Library loan statistics as a proxy for monograph selection in citation indexes», in: *Proceedings of 14th International Society of Scientometrics and Informetrics Conference* 2, Viena, Austrian Institute of Technology, 1237–1252.

Desrochers, Nadine / Laplante, Audrey / Quan-Haase (2016): «Illusions of a 'Bond': Tagging Cultural Products across Online Platforms», *Journal of Documentation* 76, 2.

Erdt, Mojisola / Nagarajan, Aarthy / Sin, Sei-Ching Joanna (2016): «Altmetrics: An Analysis of the State-of-the-Art in Measuring Research Impact on Social Media», *Scientometrics* 109:2, 1–50.

García-Carretero, Lucía / Codina, Lluis / Diaz-Noci, Javier (2016): «Herramientas e Indicadores SEO: características y aplicación al análisis de Cibermedios», *El Profesional de la Información* 25:3, 497–504.

García Rodríguez, Araceli (2013): «El papel de las plataformas de distribución y venta en la promoción del libro electrónico infantil y juvenil», *Anales de Documentación* 16, 1–19.

González Fernández-Villavicencio, Nieves (2017): «Indicadores de lectura en las redes sociales Goodreads y LibraryThing y su impacto en Amazon», *Mendeley Data* 2, <http://dx.doi.org/10.17632/ggxfsd9xx7.2>.

Halevi, Gali / Nicolas, Barnaby / Bar-Ilan, Judit (2016): «The Complexity of Measuring the Impact of Books», *Publishing Research Quarterly* 23:3, 187–200.

Kaplan, Nora (2016): «Social reading in Spain: A discourse-pragmatic perspective», *Álabe* 13, 1–21, <doi:10.15645/Alabe2016.13.7>.

Kousha, Kayvan / Thelwall, Mike (2015): «Web Indicators for Research Evaluation. Part 3: Books and Non Standard Outputs», *El Profesional de la Información* 24:6, 724–736.

— / — / Abdoli, Mahshid (2017): «Goodreads Reviews to Assess the Wider Impacts of Books», *Journal of the Association for Information Science and Technology* 68:8, 2004–2016.

Laspa, Christina (2013): «The normal distribution in LibraryThing: investigation on the combined literature of fifteen books», *Collection Building* 32, 45–153, <http://doi.org/10.1108/CB-03-2013-0013>.

Llobet Domènech, Jordi / Coma Vilaplana, Laura / Vila Boix, Núria (2016): «Xarxes socials de lectors: Un espai per treballar?», in: *14es Jornadas Catalanas de Documentación* 1, Barcelona: Col·legi Oficial de Bibliotecaris-Documentalistes de Catalunya, 1–19.

Manso Rodríguez, Ramón Alberto (2015): «¡Leer, Comentar, Compartir! El Fomento de la lectura y las tecnologías sociales», *TransInformação* 27, 9–19.

Montesi, Michela / Aragoneses, María Esteban (2014): «Does a film adaptation of a novel influence reading behaviour? The answer is on the web», *Nuovi Annali della Scuola Speciale per Archivisti e Bibliotecari* 28, 223–239.

Lluch, Gemma / Tabernero, Rosa / Calvo, Virginia (2015): «Epitextos virtuales públicos como herramientas para la difusión del libro», *El Profesional de la información* 24, 797–804, <doi:10.3145/epi.2015.nov.11>.

Richards, Anna / Sen, Barbara (2013): «An investigation into the viability of LibraryThing for promotional and user engagement purposes in libraries», *Library Hi Tech* 31:3, 493–519.

Rivera, Patricia (2014): «Redes sociales literarias: caminando hacia la lectura compartida», *Media-Tics*, <www.media-tics.com/noticia/3184/Redes-sociales/Redes-sociales-literarias:-caminando-hacia-la-lectura-compartida.html>.

Thelwall, Mike / Kousha, Kayvan (2017): «Goodreads: a social network site for book readers», *Journal of the Association for Information Science and Technology* 64, 972–983.

Torres-Salinas, Daniel / Gumpenberger, Christian / Gorraiz, Juan (2017): «PlumX as a potential tool to assess the macroscopic multidimensional impact of books», *Frontiers in Research Metrics and Analytics* 2, 1–11.

Zuccala, Alesia A. / Verleysen, Frederik T. / Cornacchia, Roberto / Engels, Tim C. E. (2015): «Altmetrics for the humanities comparing Goodreads reader ratings with citations to history books», *Aslib Journal of Information Management* 67:3, 320–336.

Nieves González-Fernández-Villavicencio, Universidad de Sevilla, Biblioteca de Económicas y Empresariales, Avda. Ramón y Cajal 32, E-41018 Sevilla, <nievesg@us.es>.