Do we need better online book review organisation?

Tjaša Jug, tjasa.jug@ff.uni-lj.si
University of Ljubljana, Faculty of Arts, Department of Library and Information Science and Book Studies, Slovenia

Maja Žumer, maja.zumer@ff.uni-lj.si
University of Ljubljana, Faculty of Arts, Department of Library and Information Science and Book Studies, Slovenia

UDK: 025.32:001.89: 004.93=111
DOI: http://dx.doi.org/10.15291/libellarium.v9i2.268
Research paper

Abstract

Introduction. Online customer reviews present one of the most important factors in book purchasing or borrowing decisions. Given that only well-organized reviews are useful, Amazon has already started linking multiple formats and editions of the same book. Nevertheless, this method is not suitable for books that have appeared in many formats and editions as some attributes do not apply to all versions.

Research questions. In our study we were interested in the aspects of a book users perceive as important and the extent to which these attributes match FRBR entities. We were also interested in the relation between specific attributes in the reviews and the numeric rating of the book.

Methods. We used content analysis on two random samples of Amazon customer reviews. The sample included a total of 600 reviews of three well-known fiction book titles that have many formats and editions that accommodate the different preferences of readers.

Results. The results show that readers take into consideration book information at various abstraction levels that match those in the FRBR model. Most reviewers comment on the book content while review readers consider reviews that comment on different aspects of a book as more helpful.

Conclusions. Given that subjective opinion is an important factor in the users’ book selection decision, it would be reasonable to rethink the presentation and organisation of online book reviews.

KEYWORDS: customer reviews, Amazon, FRBR, content analysis
Introduction

Fiction book buying and borrowing decisions depend on book characteristics, such as its content description, author or genre. Depending on the purpose and the manner of reading, one can also be interested in other information, for example the book’s appearance or the specifics of a particular edition (Pöntinen and Vakkari 2013). In addition to browsing in library and bookstore collections, such information can also be obtained from online bookstores and library catalogues. However, it has been shown that when it comes to the decision-making process, online customer reviews are one of the most trusted advertising sources (Nielsen 2015) that also affect customers’ book purchasing behaviour (Chevalier and Mayzlin 2006). Hence, the reviews should be presented and organised in a way that will satisfy diverse information needs of customers.

To optimize product information, Amazon, the biggest online bookstore, has already started linking multiple formats and editions of the same book under a single title and has, consequently, merged all customer reviews of all editions of the same book. This method somewhat resembles the Functional Requirements for Bibliographic Records, also known as FRBR. While grouping different editions together helps to reduce the number of results and offers users a better overview of what is available, merging customer reviews may cause some problems when reviewers do not focus only on the content but also comment on other aspects of the book. Therefore, one question remains: Is the current presentation of online book reviews optimal and, if not, how can we make it better? In our paper, we aim to draw attention to the need for better organisation of customer reviews in online bookstores.

Theoretical framework and literature review

Word of Mouth (WOM) has been an important information source for buying decisions even before the Internet. This kind of message has more influence on customer attitudes than market-generated information, as friends and other buyers represent a neutral filter and describe their personal experience with the products (Arndt 1967). The Internet has made access to information easier and, consequently, allowed us to meet a wide range of consumers and hear their personal opinions. While book reviews have traditionally been written by professional reviewers and publishers, today anyone can be a reviewer, regardless of their level of knowledge of books and other literature. Nevertheless, consumers read and take into account these online reviews at each stage of their purchasing process, mostly when comparing different products or when they need a confirmation that they have made the right choice (Freedman 2008). Comments that are marked as more helpful by other customers are perceived as more trustworthy and therefore have significant impact on buying decisions (Chen, Dhanasobhon and Smith 2008). Furthermore, reviews that evaluate a product with high numeric rating are more common (Chevalier and Mayzlin
2006, Fowler and Avila 2009) and have a positive effect on product sales; however, negative information usually has an even greater impact on purchasing decision (Henning-Thurau and Walsh 2003, Chevalier and Mayzlin 2006).

Online shopping and customer reviews have also changed our view of the products, because we can obtain information on their quality before purchasing or experiencing them ourselves. Nelson (1970, 1974) distinguishes between search and experience goods. Search goods are those products that consumers can seek, inspect and compare opinions on prior to purchasing them and thus familiarize themselves with their quality before making a purchase. In contrast, consumers cannot be informed of the quality of the experience goods in advance as they need to use them for some time after the purchase to evaluate them (Nelson 1970). According to this categorization, the book content can be considered an experience good as readers actually need to read it to determine whether they like its style, dynamics, etc.

On the other hand, Klein (1998) posits that the Web can transform experience goods into search goods, since a variety of databases and tools have made search for any product characteristics much easier and less costly. This has also changed our way of evaluating different characteristics and enabled us to have a virtual experience of a product via demo versions, previews, and through the eyes of other customers (Klein 1998). From this point of view books have transformed from experience goods into search goods or even into a hybrid between experience and search goods (Nakayama, Sutcliffe and Wan 2010). Some researchers (Bae and Lee 2011, Huang, Lurie and Mitra 2009) have also found that reviews and their effect on perceived helpfulness differ across product types. In terms of buying decisions, customer reviews are a more important source for experience goods than for search goods. Furthermore, extreme positive ratings in reviews are more important for search goods, while in-depth reviews with moderate ratings are more helpful for experience goods (Mudambi and Schuff 2010). It is important to note that book buyers have different needs and are not always interested only in the book content, which is an experience attribute. They can also be interested in the characteristics of a particular edition, different formats of the same work, or in augmentations, such as prefaces, author’s biography or study guides.

Various research results report on a positive impact of online reviews on the book purchasing decision (Chevalier and Mayzlin 2006, Lin, Huang and Yang 2007) and also on the users’ borrowing intention (Huang and Yang 2010). Researchers (Kakali 2014, Pecoskie Spiteri and Tarulli 2014) have found that user generated content such as online reviews and folksonomies represent an important subjective addition to the professional book description in library catalogues as they often provide subjective information, such as book’s dynamics, emotional reading experience, similar titles, etc. They therefore suggest cooperation between libraries and online bookstores and see customer reviews as an additional promotional channel (Huang and Yang 2010, Kakali 2014). The current use of
library catalogues for finding and consequently selecting fiction is relatively low (Mikkonen and Vakari 2012) because they lack the needed data and are difficult to use (Švab, Mercun and Žumer 2014). Some libraries have already integrated in their online catalogue various social discovery platforms, such as Bibliocommons (http://www.bibliocommons.com/), Encore Discovery Solution (http://www.iii.com/products/sierra/encore), and SirsiDynix (http://www.sirsidynix.com/). These social catalogues allow users to connect with each other through user generated content and are much more similar to online bookstores (Pecoskie, Spiteri and Tarulli 2014). Given that the integration of customer reviews in library catalogues has not yet become a common practice, readers still often seek subjective opinion about books from other sources, such as online bookstores and other book-related platforms.

Amazon, the biggest online bookstore, offers more than 65 million books and has a very strong community of product reviewers. To improve the organisation and overview of its vast collection, Amazon is grouping all formats and editions of the same book under a single title and, consequently, all customer reviews of that book. This view somewhat resembles the Functional Requirements for Bibliographic Records (FRBR), a library-based conceptual entity-relationship model (Functional… 1998). FRBR considers a book at various levels of abstraction, where Work represents the abstract idea, which is realized through one or more Expressions - another abstract entity. Manifestation is the physical embodiment of one or more Expressions, while any exemplar of a manifestation is called an Item. Therefore, a book as a product is not represented only by its content but also by other characteristics related to its format or edition. FRBR is a user-centred model and it is not useful only in library catalogues, but could also find its place in online bookstores, since it supports user tasks, such as the finding, identifying, selecting and obtaining of the desired entity. Its intuitiveness and capability to take advantage of new technologies support meaningful grouping of the results, browsing, exploring and identifying the relations among entities (Functional… 1998).

Grouping reviews of all formats under a single title is often convenient, especially when readers are interested in the content. The problem emerges with widely known books that have many formats and editions and therefore usually a lot of customer reviews. Too many reviews can discourage customers’ desire to purchase a product (Lin, Huang and Yang 2007), thus many online book platforms have integrated various sorting and filtering tools to reduce the number of results. Despite a smaller and better set of filtered reviews, users typically read only the first few. Their opinion of a product is therefore formed on a small sample of reviews, which do not necessarily include relevant information (Liu, Karahanna and Watson 2011). We can assume that buyers might be interested in various aspects of a book and that merging customer reviews may not always be appropriate, especially in situations where a book has appeared in several versions. What is more, users can encounter reviews with inconsistent textual and numeric ratings of a product, which can result in
a bad shopping experience. This phenomenon occurs due to various reasons, such as individual interpretation of ratings or a mixture of positive and negative opinions about different product attributes, and is more common for reviews of experience goods (Mudambi, Schuff and Zhang 2014).

Some researchers (Liu, Karahanna and Watson 2011, Huang et al. 2014) believe that it is necessary to improve the current online customer review presentation. The idea is to categorize online reviews by meaningful product attributes and thus help consumers to better comprehend and interpret reviews (Liu, Karahanna and Watson 2011). In this way, users would be able to identify all the associated attributes of the product and read only those reviews that include information about attributes they are interested in. Although identifying attributes is much easier for search goods, it could also be done for experience goods and could therefore also be useful for organising book reviews which have features of both products types.

**Research questions**

Customer reviews are a very important factor in the purchasing decision, but only if they are well-presented. When there are too many reviews, readers get an incomplete picture of the product and are less likely to find the information they are seeking. Although review filtering tools and numeric ratings can be very helpful, they can become useless or even misleading when we are not sure whether the reviewers rated the content, book's features or their shopping experience. Some researchers have already proposed (Liu, Karahanna and Watson 2011) review categorization by product attributes. Despite the fact that readers are usually interested in the book content, which is an experience good, they can also be interested in its format, edition, version, etc. By grouping reviews, Amazon has satisfied the information needs of the first group, but this does not predict the needs of the latter.

In our study, we have examined online book reviews from two points of view. Firstly, we were interested in the aspects of the book reviewers refer to when writing their comments and the extent to which these attributes match FRBR entities. We were also interested in the relation between specific attributes in the reviews and the numeric rating of the book. Secondly, we wanted to determine the type of reviews that seem to be useful to review readers as well as the frequency and type of entities mentioned in therein. If it were to be determined that the majority of the reviews focus on the book content, i.e. the Work, it would be reasonable to conclude that grouping reviews of different book versions is useful. On the other hand, if it were to be determined that reviews include information on different aspects of a specific book version, it would be reasonable to rethink their organization and the possibilities of their presentation.
Methods

The study was conducted in March 2015. Content analysis was conducted on two samples of Amazon customer book reviews, with an aim to determine the presence of the FRBR entities in the online book reviews. First, we selected three well-known book titles that have many formats and editions and attract different types of readers (The Little Prince by Antoine de Saint-Exupéry, Adventures of Huckleberry Finn by Mark Twain and Frankenstein by Mary Shelley). These classics also have movie adaptations and are among the topics covered by school syllabus, which means that they are read for various motivations. Our criterion was also that they had a similar number of reviews at the time of our study.

A random sample of 100 customer reviews for each title (300 in total) was selected. We have found that the percentage of sampled reviews for each star rating was very similar to the actual number in the population. Similar to the findings of other researchers (Chevalier and Mayzlin 2006, Fowler and Avila 2009), the majority of reviews (85 %) were rated as good (4 and 5 star ratings), while only 15 % received low ratings (3, 2, and 1 star ratings). This sample represents reviewers who have already had an experience with the identified book, and an opinion on important aspects of the book that need to be outlined.

Additionally, for each title, a sample of 20 reviews marked as most helpful for each star rating (300 in total) was selected. This sample included reviews that the review readers found most helpful and moved to the top of the list by posting comments. These comments were visible to the majority of users as we assumed that they read up to two pages of reviews. Our sampling process is shown in Figure 1.

Figure 1. The sampling process and the number of reviews
Before carrying out the content analysis, we created a coding scheme, which was tested with a pilot study on Amazon reviews of five different book titles. The scheme was designed to determine the presence of FRBR entities in the online book reviews. For example, work-level opinions would comment on the content, genre and author, while expression-level reviews would mention the quality of a translation, an abridgement or some other changes to the text. Reviewers may also talk about the manifestation by remarking on the format, illustrations, typography, etc. or about the item when they comment on the condition of a specific copy. We also examined whether the reviewers mentioned different expression, different manifestation, related work, unrelated work or item/shopping experience. We marked all the identified entities in reviews, regardless of their frequency or the focus of the specific review. Each entity was assigned only once for an individual review.

Figure 2 shows an example of our entity identification process on a review of the Adventures of Huckleberry Finn. In this case we identified the entity Work where the reviewer expressed his opinion on the content of the book and the entity manifestation where he complained about the appearance of the text and the quality of paper of this version.

This is also a good example of a review with inconsistent overall numeric and textual rating of the book. While a reviewer likes the book’s content, he/she is not satisfied with the physical appearance of a particular edition, which he/she gave the lowest rating. A different type of buyer might give the same version an average rating, which would reflect his preference towards the content and a dislike of the appearance, or a high star rating while he would highlight the book’s weaknesses in the comments section. This could affect users’ opinion about the book as they do not know which characteristics have affected the overall rating of the title. This was evident in 2010 when Lewis’ book The Big Short was released. The book reached the number one sales rank spot on Amazon, but had very low numeric rating. There were more one star ratings than 2 – 5 star ratings combined; however, more than half of the one star ratings did not comment on the content. Instead, most of the reviewers expressed their dissatisfaction with the fact that the Kindle version had not been released at the same time as the hardback edition (Carr 2010).
Research results

The results show that overall there are more entities identified in the sample of readers than in the sample of reviewers. 474 entities have been identified in the sample of reviewers and 696 entities in the sample of review readers. This was to be expected knowing that more detailed reviews offering opinions on many different aspects are more likely to satisfy the readers’ information needs.

The majority of reviews consist of a smaller number of entities, especially in the sample of the reviewers, where more than half of reviews comment on only one aspect of the book. On the other hand, the majority (32 %) of reviews in the readers sample refer to two book aspects. There is a greater number of overall entities identified, which indicates that review readers perceive comments that include more entities as more helpful. Figure 3 shows that the number of reviews declines as the number of entities they include grows.

![Figure 3. The number of reviews containing different number of elements](image)

If we look at individual entities in the reviews we can get a better overview of the most frequently commented aspects of the book and a better insight into information that the review readers find important. We have found that 261 reviews (87 %) in the sample of the reviewers and 212 reviews (71 %) in sample of the review readers contain the entity **Work**, which is, as expected, the most frequently represented entity in book reviews. Other common entities in the first sample are **Related work** (51), **Manifestation** (40), **Expression** (37) and **Item** (29). The results are different in the second sample, where the aspects of the described books are more important than references to other works. The second most frequent entity in this sample is the **Expression** (123) which is followed by **Different expression** (96), **Manifestation** (94), **Related work** (76), **Item** (38) and **Different manifestation** (27). It is interesting that **Work** is the only entity that appears less frequently in the sample of the review readers than in the sample of the reviewers, which could indicate that reviewers often mention the book’s content, while buyers are more interested in other aspects of the book. These results could be related to the fact that we have intentionally studied reviews of the books that are well-known and have appeared in multiple versions. We
can also notice that *Related Work* is mentioned quite frequently in both samples. This is also the result of the selection of the titles for our research. In other words, reviewers often compared book and movie adaptations of *Frankenstein*, or mentioned the book *Tom Sawyer*, which is related to the *Adventures of Huckleberry Finn*.

We also looked at the results from another angle and examined the frequencies and combinations of different entities mentioned in the reviews. We have found that 58% of the reviewers in the first sample commented only on the *Work*-level of the book or used the *Work/Unrelated work* combination, i.e. entities suitable for describing all book versions. The percentage of the reviews which also refer to other book attributes is bigger in the second group as the *Work*-related comments account for only 29% of the sample. The proportion of such reviews in the case of the *Adventures of Huckleberry Finn* is quite large even in this sample because reviewers often comment on the controversy of the content, which also seems to be important for the buyers of this title.

**Figure 4. Proportion of Work-related reviews of a specific title and the percentage of reviews containing other entities**

Furthermore, we have found that 13% of the reviews do not mention the *Work*-level attributes in the first sample. This percentage is even bigger in the second sample, where 29% of the examined reviews do not comment on this abstraction level. An equal percentage of reviews describe the book exclusively on this level. This piece of information is even more important given the fact that reviewers often briefly comment on the content although they mainly want to emphasize other characteristics of the book. In the sample of the review readers, the *Work/Expression/Different expression* combination (25) and *Expression/Different expression* combination (13) are also common. This pattern is mostly evident in the case of *The Little Prince*, where a lot of reviewers compared two different translations of the text. There were also 7 cases where we have identified the presence of four different aspects of the book (*Work/ Expression/ Different expression/Manifestation*). In 9 reviews users also commented on
the *Manifestation* and their experience with the *Item*, which is a combination corresponding to a specific version of the book or even a unique copy.

As we have noted before, the review process can be inconsistent especially when reviewers want to describe their experience with different characteristics of the book. Therefore, an overall book rating could be misleading not only on online platforms, where reviews are merged under a single title, but also in any other system that does not allow users to separately rate different aspects of the book. This is evident in our sample of the review readers, where we have examined the distribution of entities present in the comments for each numeric star rating. In Figure 4 we can once again see that most reviews describe the book on the *Work* level. We assume that most of these textual reviews are very positive as they are mainly given 5, 4 or 3 star ratings. The diagram on the left shows the distribution of the reviews that contain *Manifestation* and *Different manifestation*. We can see that reviewers who commented on the book characteristics on this abstraction level assigned extremely high or low numeric ratings. This could indicate that buyers feel the need to express their opinion about the physical appearance of a book only when they are very satisfied or dissatisfied with a particular version. A similar pattern was observed in the reviews that talk about *Different manifestation* as reviewers often compare it to the commented version. There is also a similarity in the pattern observed in the reviews that comment on the *Expression*, *Different expression* and *Item*. The reviews commenting on a specific *Expression* are typically negative and often remark on weaknesses of the text version such as the translation, and compare it to another version, i.e. *Different expression*. It also appears that buyers report on the *Item* or their purchase only when they have had a bad shopping experience.

![Figure 5. Number of entities in reviews marked by a specific rating](image)

These findings somewhat concur with the conclusions drawn by Mudambi and Schuff (2010), who claim that extremely good ratings are more important for search goods, while in-depth reviews with moderate ratings are more helpful for experience goods. In this case, the content as an experience good was present in most of the reviews marked as helpful, while search goods have typically been...
given extreme numerical ratings. Therefore we can also agree with Nakayama, Sutcliffe and Wan (2010), who see books in the Internet age as a hybrid between search and experience goods.

Discussion

In our study, we have found that Amazon’s method of grouping different book reviews under a single title could be effective if they concerned only the Work-level aspects of the book. Although reviewers most frequently express their opinions on the book content, review readers often perceive reviews that comment on various other book attributes as more helpful, especially in the case of well-known titles which have many formats and editions. The percentage of the most helpful reviews that comment only on the Work-related aspects of the book is equal to the percentage of the reviews that do not even mention any book attributes on this level and therefore do not comment on the book content.

What is more, reviewers most often rate a book as good when commenting on the content, while the negative ratings frequently refer to other aspects and lead to a misleading overall rating of the book. From this we can conclude that merging book reviews is suitable for the titles with fewer versions while it is not efficient for the titles that have appeared in many formats and editions.

Considering that online reviews are one of the most important factors in the book buying and borrowing decision, it would be reasonable to rethink their presentation and organisation. As proposed by Liu, Karahanna and Watson (2011) reviews could be grouped into categories by product attributes described therein, where each attribute could be given a separate numeric rating. This could be achieved in three different ways. The first option is an automatic computer classification which requires minimum human interaction, but is the least accurate and reliable of all. The second option could be giving reviewers an opportunity to comment on different aspects of a book, such as its content, writing style, typography, illustrations, complexity, etc. This method is already used in some online stores for search goods but it is not very suitable for reviewing books as it affects the natural mind flow and solicits reviewers’ comments on aspects they are not interested in. In our opinion the optimal way of organizing book reviews would be to make it possible for users to classify reviews using social tagging as they are the ones who know best what they need and which book attributes help them decide whether to buy/borrow a specific book. Furthermore, book reviews could be categorized, which would enable users to filter and read only those containing the attributes they are interested in. As free tagging has some weaknesses, such as too many low-frequency terms with similar meaning, it would be reasonable to determine the categories and offer only a few that represent different aspects of a book, as those represented in the FRBR model.
Conclusion

Online bookstores and other platforms that support user-generated content are a common source of book information. Although readers are usually interested in the book content, some of them are also interested in other aspects, which could be crucial for their final buying or borrowing decision. This is an important finding for the presentation and organization of book information in an online environment, not only in bookstores but also in library catalogues where subjective opinion about book attributes could represent an addition to the professional book description. In our study, we used content analysis of Amazon customer reviews and found that merging reviews is not very efficient for books that have many formats and editions. However, this does not necessarily mean that it is not useful for the titles that have appeared in fewer versions or do not have such diverse audience. Nonetheless, given that online book reviews are an important source of information for customers' book buying/borrowing decision, it would be reasonable to rethink their organization and the possibilities of their presentation. This could be done by grouping customer reviews by product attributes described therein. Given that the results show that people consider book information at various abstraction levels that match those in the FRBR, this model could serve as the basis for the creation of meaningful categories that would enable individuals to filter and select only those reviews that describe a particular aspect of a book. In the future, similar studies could be done on a bigger sample and with different types of books, including nonfiction titles. In addition, further research could be conducted involving book buyers who could test the new presentation and compare it to the current organization of online book reviews.
References


