Content, Physical Appearance, Copy Condition

Tagging Customer Book Reviews

Tjaša Jug, Maja Žumer

University of Ljubljana, Faculty of Arts, Slovenia
{Tjasa.Jug, Maja.Zumer}@ff.uni-lj.si

Abstract

Users of online bookstores are not interested only in general book description when searching and buying books, but also in subjective reader opinion, which could be found in online reviews. Reviewers usually comment on book content, but may also mention other aspects of the received book, such as binding, illustrations, translation etc. Meanwhile buyers are not always interested in the same aspect of a book, especially when they need it for a special purpose. Currently, obtaining non-content book information from reviews is difficult, therefore it would be reasonable to rethink their presentation and organization. In our study, we used an interview and a task solving method to determine whether social tagging could be an appropriate aid for this purpose. The results show that free tagging offers insight into users' vocabulary but is not optimal for online review presentation. Nevertheless, it represents a good basis for creation of categories that describe books on different levels of abstraction and could be used as a filtering tool, which would select only those reviews containing the aspects of a book a buyer is interested in.

Keywords: books; reviews; tagging; Amazon; book aspects; FRBR

1   Theoretical framework

Customer reviews are often integrated in online bookstores and other book-related platforms as they represent an important factor for book purchasing and borrowing decision (Chevalier & Mayzlin, 2006; Lin, Huang & Yang, 2007; Huang & Yang, 2010). Some online bookstores, such as Amazon, have made a step further by combining all formats and editions of a work under a single title and all customer reviews for a work. This method somewhat resembles the Functional Requirements for Bibliographic Records (FRBR) (Functional …, 1998) which considers a book at various levels of abstraction (*Work, Expression, Manifestation* and *Item*). Book reviews do not focus only on the *Work*-level aspects of a book, such as content, genre, author etc., but also on the characteristics of a particular edition. What is more, they can also comment on the received copy (an *Item*) or their buying experience (Jug & Žumer, 2016). Meanwhile, book buyers might be interested in various aspects of a book not only when they are searching for a book for a specific purpose, such as school reading assignment or a gift, but also when they can choose among versions with different translators, illustrations, text design or binding. Therefore, merging customer reviews may cause problems when reviewers express their opinion on aspects of a book that do not correspond to all versions of the book. Some researchers (Liu, Karahanna & Watson, 2011) have already found that it would be reasonable to group customer reviews into categories according to product attributes described in the reviews. Due customer interests in various book attributes, also online bookstores could use a filtering tool that would help individuals to select only comments describing a particular aspect of a book. As users know best what they need, it might be useful if they could classify reviews with social tagging and create categories for different book characteristics. We are assuming that user generated content could be helpful for this manner as it reflects their interests and vocabulary (Spiteri, 2007).

2   Research questions and methods

In our study, we aimed to identify which aspects of a book people recognize while reading book reviews and how do this aspects overlap with the FRBR

entities. The FRBR model was chosen because it was proven to be intuitive (Žumer, Salaba & Zhang, 2012) and it provides an appropriate clustering mechanism. We were also interested in vocabulary people use to describe different book characteristics recognized in the comments and whether a social tagging would be an appropriate method for different review organization.

Prior to the study, we modified 11 Amazon reviews of 9 well known book titles with many versions in a way that they contain opinions on diverse book aspects and include different combinations of book attributes. The most of the reviews commented on printed books, but there were also two comments for audio books and one for a kindle version. After preparing the reviews, we carried out a content analysis, where we determined which FRBR entities match attributes mentioned in our comments. In the second step of our study, twenty-five people read reviews and added their own tags to describe the book attributes they had identified in the reviews. After the data collection, we compared obtained tags to the FRBR entities and designed 10 categories in a way that covers all four levels of abstraction of the FRBR model and at the same time reflects participants' terminology. Names of the categories were based on the most frequent words assigned by participants to the part of the text describing the book on the particular level of abstraction. For example, category *Content* applies to FRBR's *Work*, *Edition* to *Expression* or *Manifestation, Medium, physical appearance* and *Additions* to *Manifestation* and *Copy condition* or *Shopping experience* to *Item*. We also added category *Target audience* and categories *Comparison with different editions* and *Comparison with different works*, which describe related products.

One month after the first round, we asked the same participants to read the same reviews and to identify book aspects mentioned in them by marking relevant categories from the list. We also tested the applicability of these categories with the same set of reviews and a new group of twenty-five participants in round three and compared the detected aspects in three different rounds. During the study we did not offer the participants any other information about the categories, as we wanted to test how self-explanatory are they.

Figure 1 represents a review for the special edition of The Hobbit. In this review, we identified FRBR entity *Work* where the reviewer mentions the book's content. While he likes the cover, binding and illustrations, which are attributes of a *Manifestation,* his disappointment and rating refer to the received copy – an *Item*. He also mentions *Related work* (Lord of the Rings).

197 of 216 people found this helpful
★★☆☆☆ **Poor quality**
By Collector; 21. November, 2004
Format: Hardcover

Note: Tolkien's work is rated 5 stars. The physical book is rated 2 stars.

I am a big fan of Lord of the Rings and I read this book many times. When you pay $100 for a collector's edition book, you expect it to be perfect in all respects. The cover, binding and illustrations are very beautiful, but there are many pages glued together in my copy. Some are even wrinkled or torn. I am very disappointed.

Fig. 1 Example of a review used in our study

In figure 2 one can see some of the tags and categories assigned to the review from figure 1 by participants in two rounds. For this review, first group of the participants assigned 78 unique tags in round 1. As we can see in the middle column, we obtained many tags with similar meanings and low frequencies, such as *Collector's edition, Special edition, Physical edition* etc., which would not be appropriate for automatic processing of book reviews. As seen in the right column, choosing from a predefined list in round 2 proved to be a better approach as these categories are easy to select and are therefore more transparently reflecting the most noticeable and important aspects of a book in a specific review.

FRBR	FREE TAGS	CATEGORIES
Work (15)	**Content (4)** **Work evaluation (2)**	
Manifestation (16)	**Collector's edition (4)** **Physical appearance (2)** Special edition (1) Physical edition (1) Concrete edition (1)	**Edition (14)** **Physical appearance (18)**
Item (23)	**Glued pages (3)** **Evaluation (2)** Condition (1) Product quality (1) Bad shopping experience (1)	**Copy condition (22)** **Shopping experience (16)**
Related work (3)	Other work (1) Continued by (1)	
	Personal opinion (3)	

Fig. 2 FRBR entities, tags and categories added in both rounds to the review from figure 1 (number of occurrences)

3   Research results

After each round, we asked the participants to determine the difficulty of the task and whether they would use the tagging system if it were integrated in the online bookstore. Figure 3 shows their responses. The most participants from the first round consider free tagging as moderate or difficult (22) and they said that they would not tag customer reviews in this way (16). On the contrary, they said the task was easy (14) when they had the list of categories and that they would use this tool (16), especially for the books they are interested in. The results were similar in the third group where the participants did not use free tagging method.
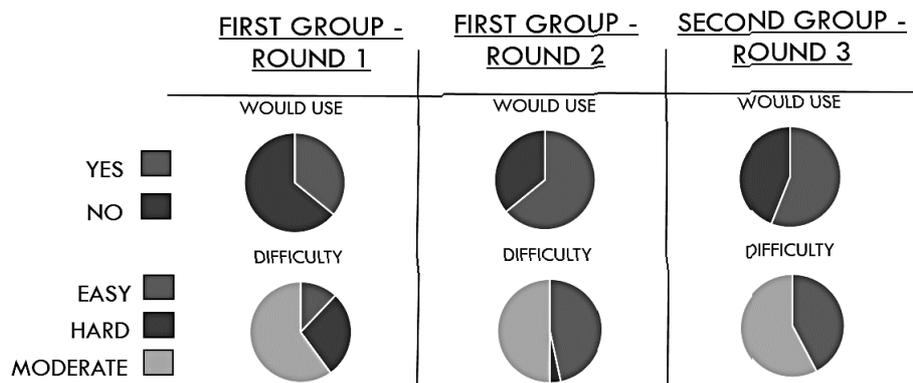

Fig. 3 Use of tags and difficulty of the task

The participants were also asked to comment on the pre-made list of the categories. Most were satisfied with the number of categories, but they had problems understanding some of them. They were not sure which book information falls under the term *Additions*. They also said that they are unsure about categories *Target audience* and *Medium*, although the results show that they chose these categories properly. Some participants also did not distinguish between categories *Physical appearance* and *Copy condition* and therefore often marked them both. They also suggested some new categories, for example *Subjective opinion, Value, Comparison with different medium, Technical details* and a *Free category,* where they could add their own tag.

4   Conclusion

The results indicate that people consider book information at various abstraction levels that match those expressed in the FRBR model. We also found that free tagging could be an effective method for identification of book attributes mentioned in the reviews but is not optimal for their organization and presentation. Meanwhile, a controlled list of categories that represent different aspects of a book could be a useful tool for this purpose and would enable automatic grouping and filtering of reviews. However, we cannot be sure if this technique would actually be used in an online bookstore context, therefore more research on a bigger sample and on an actual implementation of this tool would be needed.

References

Chevalier, J. A. and D. Mayzlin (2006): The effect of word of mouth on sales: Online book reviews. In: *Journal of marketing research*, 43 (3), 345–354.

Functional Requirements for Bibliographic Records: Final report (2009). Munich: K.G. Saur Verlag.

Huang, Y. and W. I. Yang (2010): A study of internet book reviews and borrowing intention. In: *Library Review*, 59 (7), 512–521.

Jug, T. and M. Žumer (2016): Do We Need Better Online Book Review Organisation? In: *Libraries in the digital age (LIDA)*, Zadar.

Lin, T. M., Y. K. Huang and W. I. Yang (2007): An experimental design approach to investigating the relationship between Internet book reviews and purchase intention. In: *Library & Information Science Research*, 29 (3), 397–415.

Liu, Q. B., E. Karahanna, and R. T. Watson (2011): Unveiling user-generated content: Designing websites to best present customer reviews. In: *Business Horizons*, 54 (3), 231–240.

Spiteri, L. F. (2007): The structure and form of folksonomy tags: The road to the public library catalog. In: *Information technology and libraries*, 26 (3).

Žumer, M., A. Salaba, Y. Zhang (2012): User Verification of the FRBR Conceptual Model and Testing of FRBR Prototypes. In: *International Conference on Asian Digital Libraries* (pp. 148–151). Berlin: Springer.