Researching bibliographic data with users: examples of 5 qualitative studies

Katarina Švab
University of Ljubljana, Faculty of Arts, Department of Library and Information Science and Book Studies, Slovenia. Email: katarina.svab@ff.uni-lj.si.

Tanja Merčun
University of Ljubljana, Faculty of Arts, Department of Library and Information Science and Book Studies, Slovenia. Email: tanja.mercun@ff.uni-lj.si.

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

Abstract
Introduction. Library catalogues enable people to explore and take advantage of the wealth of library collections. However, their use is relatively low, not only because they are difficult to use but also because they lack the needed data.

Research questions. To go beyond the constraints of current bibliographic data and find potentially missing data elements, our research investigated what data is needed to help different types of users find, identify, select, obtain, and explore information in the context of fiction.

Methods. Using a combination of qualitative methods (observations, surveys, and interviews), different groups of users were investigated. For each of the groups a special study was designed to find out based on which criteria they selected books. Rounding up the series of studies, a focus group and interviews were organised with reference librarians to tap into their rich experience.

Results. Although the paper briefly outlines some of the main conclusions from the five studies, more focus is given on the study descriptions from the viewpoint of their design.

Conclusions. To improve digital or classical services, investigation of information needs is one of the key areas that can benefit considerably from qualitative research methods. Our paper provides examples of how these studies can be designed and what kind of research questions they can help us answer.

Keywords: information needs, interviews, observations, focus groups, library catalogues

Introduction
Library catalogues present the central tool that enables people to explore and take advantage of the wealth of library collections. However, their use is relatively low, not only because they are inefficient and too complicated (Calhoun et al., 2009) but also because they lack the needed data (Hypén, 2014) that would help users as well as librarians not only find, but also select, identify, and explore the desired materials. This indicates that changes are needed if libraries wish to provide valuable services and make the best use of their collections. What has often been forgotten is that it is not enough to only build more modern systems, it is essential that they are centred around users’ needs and the information seeking process.

While fiction represents an important part of (public) library collections and circulation, its retrieval presents one of the major problems in current catalogues also because it often leads to long lists of results where it is difficult to distinguish between different editions of the same work or explore its various versions.

Studies (Mikkonen & Vakkari, 2012; Goodall, 1989, Pogorelec, 2004) show that only between 10 and 20 percent of adult readers use the library catalogue to access fiction. With the catalogue predominantly supporting only known-item searches, users have developed tactics for finding good fiction books without the help of library catalogue (Ross,
2001) by scanning or browsing bookshelves, an approach that is becoming more difficult as collections grow in size (Hypén, 2014) and as more and more books becomes available in electronic form.

Current online library catalogues are thus faced with two tasks connected to fiction: a) to support tactics other than known-item search (Saarinen & Vakkari, 2013) and b) to provide all the needed data for identification and selection of fiction based on bibliographic description. Pöntinen & Vakkari (2013) point out that especially with the rise of e-book collections, it is necessary to study how readers select books by using metadata in order to inform the design of metadata for fiction. This is also true for traditional collections as users want to be able to determine a book’s relevance using their computer and expect information to assist them in this evaluation (Calhoun et al., 2009). All this suggests that libraries need to make it easier for users to determine whether the items meet their needs without examining the physical copy (Chercourt & Marshall, 2013). Also Saarinen & Vakkari (2013) observe that there is a lack of studies analysing from which attributes users infer that the book is what they are looking for. Another interesting question that arises is also whether and how different types of library users view bibliographic data in the retrieval process (Tosaka & Weng, 2010).

Our research therefore set out to investigate how people select fiction based on bibliographic records and how in physical form, thus trying to elicit what bibliographic data is needed to help different types of users find, identify, select, obtain, and explore information in the context of fiction. Are all the decisive elements presented to the users or should library catalogues be enriched with additional information?

**Literature review**

Research on enriching bibliographic records in library catalogues has a long history (for example Cochrane & Markey, 1983; Matthews, 1983). When asked what additional features users would wish to see in a catalogue entry, they most commonly requested summary, abstract, and other content information. Also more recent studies (Calhoun et al., 2009) found that tables of contents and abstract/summaries are among the most desired data-quality enhancements for end-users. Not only significant from an informative point of view, a number of research (Dinkins & Kirkland, 2006; Morris, 2001) shows that enriched bibliographic records have an important influence on circulation. Chercourt and Maschall (2013), for example, report that there is a positive correlation between adding tables of contents and increased circulation for certain groups of items, especially older materials.

Information about some important bibliographic elements can also be found in studies that investigate how people select books they wish to read for pleasure. Ross (2001), for example, reports on 194 intensive open-ended interviews with adult readers which, among other, reveal that author, genre, cover, title, sample page, and publisher give readers important clues on the reading experience they can expect and that subject, setting, and the physical size of the book help them match their book choices to the desired reading experience. Similarly Saarinen & Vakkari (2013) looked at which attributes readers perceive as indicators of a good novel and what tactics they use to find such a book in a public library. Using observation and semi-structured interviews with 16 adult library users, one of their main conclusions was that systems should support fiction retrieval by reader typology.

Focusing on children’s literature, Anderson et al. (2001) studied how parents selected books for their four-year-olds. 12 fathers and 12 mothers were asked to pick out 5 out of 14 books to read to their children in the following week and to give reasons for their selection. While the choice was somewhat dependant on the parent’s gender and the gender of the child, the reoccurring criteria were also the aesthetics of the book, familiarity with the particular book or author, content, educational value, reading level, values, children’s interests, and general trends.

Pöntinen & Vakkari (2013) analysed how readers select fiction in online public library catalogues and compared whether there are differences in the selection between an enriched catalogue and a traditional one. Using eye-tracking, 30 participants were tested in a between-subject experiment where the researchers examined which elements were most important to users by following their gaze. In contrary to some other studies, they observed that users’ choice was based on external attributes of books, whereas the content description did not seem to be as crucial.

Using think-aloud sessions, Hoder and Liu (2013) asked 20 participants to complete 10 tasks in a library catalogue and verbalize their thoughts, specifically those relating to their use of record elements. The study showed that participants found title, author, subject, year, material type, edition, table of contents, and co-author most useful.

Chang (2012) also investigated which key points help students make a decision. Carrying out interviews and observations with 60 students, she found out students use enhanced bibliographic elements for selection and identification of needed resources and that “excerpted contents” and covers helped them make a decision between different available versions.

Besides using observations and interviews with users, some researchers also applied content analysis methods to identify attributes used not only in library catalogues but also in other services such as social cataloguing sites and online bookstores (for example Adkins & Bossaller, 2007; Šauperl, 2013). The results indicate how bibliographic records could be enriched and what information and services users might expect from the library (Spiteri, 2009).

An interesting aspect has been researched also by Pejtersen (1977), who looked at library catalogues from the
perspective of reference librarians. She argued that librarians faced two main difficulties: first, the problem of identifying the user’s needs, and then the problem of formulating a relevant search strategy among books which are not classified according to the needs and which classification characterizes only some aspects of the book, insufficient in the view of the multi-dimensional needs of the user. Analysing almost 300 user-librarian conversations, the author identified five main dimensions of user needs: subject matter, frame (time, place), author’s intention/attitude, accessibility (readability, physical characteristics such as typography, modern/old, series, size, and volume), and other formulations (author’s name, title, similar books etc.).

Research agenda: 5 studies

Many of the studies looked at more traditional elements in bibliographic records with the addition of cover and some added content summaries, but did not try to introduce other information that is not part of current cataloguing practice. Our research wished to investigate more closely a wider range of elements (not only related to content, but also to attributes) that might be important to users when they search and select fiction books. To do this, we designed studies with both users and librarians who answer users’ requests on daily basis.

3 different groups of users (mothers of pre-school children, high-school students and adults looking for leisure reading) were each given a set of tasks where they operated with both bibliographic records and physical copies of the books. Observing their decisions and questioning them on how they made their choices or why they changed their decision on the book they selected enabled us to get a closer look at which elements presented the key factors as well as which elements might not be included in the current records, but were obviously important. We were also interested in how these key elements varied among different user groups and whether enriched records improved the users’ satisfaction with the chosen book. Besides examining users, we also felt that librarians would be able to provide a good insight into the topic, which is why we designed a focus group and an interview study to tap into their experiences.

Mothers of pre-school children

_Aim._ To establish whether parents are able to find a suitable book and differentiate between different texts and editions bearing the same title solely using the information recorded in a bibliographic record.

_Data collection technique._ Questionnaire and observation.

_Research questions._ Are current bibliographic records appropriate for the selection of books for small children? Deciding among several similar books, how do parents make their selection when using bibliographic records in a library catalogue and how when they choose between physical copies in a library?

Study Design. Six bibliographic records found under a title search “Cinderella” were selected and printed from an existing library catalogue. Issued in different years, in varying sizes, with different illustrations/translator, and even as adaptations, such a set of records presented a realistic search result in a library catalogue that any user searching for a story of Cinderella would have to handle.

_Procedure._ 26 mothers of pre-school children (under 6 years old) with varying levels of education were included in the study. The interviews took place in June and July 2011 outside the library setting: at children’s playgrounds, in the parks, on the beach etc.

After some general questions about picture books, libraries, bookstores, and library catalogues, six bibliographic records were presented to the mothers. They were asked to choose a record they found most suitable for their child and to comment which attributes the decision was based on. Afterwards they were presented with the book described in the chosen record and requested to comment on their satisfaction or dissatisfaction with the book. In the end, participants were handed all six books with a question which of them they would choose as most appropriate for their needs and would therefore hypothetically wish to borrow.

Results. In the Cinderella bibliographic records, parents paid most attention to the author, the publication year, form of work, the translator, and the extent of the book. However, when presented with the book they selected using bibliographic records 19 out of 26 mothers were not satisfied, the main argument being illustrations and the physical condition of the copy. After seeing all the available versions in a physical form, as many as 22 mothers said that they would prefer a different version from the one they got based on bibliographic records, the main reasons being illustrations, original text, and the condition of the copy.

1. In general it seemed that illustrations, content, the cover, the size of the book and the length of the story are most important when parents select picture books for their children. 21 out of 26 mothers said that it is usually important to them which version or edition they borrow.

2. Interviews revealed that elements such as paper thickness (for cardboard books), cover image, sample pages (as in the case of some online bookshop catalogues, such as Amazon), font size, letter case, page layout, preservation, age appropriateness would be welcome in a library catalogue.

_Comments._

1. The study was not carried out in a library setting, thus including also mothers who do not visit the library.

High school students

_Aim._ To get an overview of bibliographic data that is important to young students when they select books for their home-reading assignments.

_Data collection technique._ Questionnaire and observation.
Study Design. 11 editions of Don Quixote, a work listed as required reading in high school, were selected (abridged editions, full editions in two or four volumes, different translations and additional contents such as forewords and biographies etc.). The bibliographic descriptions were not taken from existing catalogues but were created (content and form-wise) by ourselves, using traditional bibliographic elements presented in catalogues as well as some attributes and relationships of our own choice (weight, binding, colour of the cover, short description, contents etc.). All records were presented in a mindmap where the elements were logically grouped.

Procedure. 105 high school students from two different secondary schools were included in the study which took place in November 2010. Students were recruited in the school library and on the hallways during breaks.

Presented with 11 bibliographic records, students were asked to select the one that would best fit the needs of their home reading assignments. The chosen book was then handed to the participant who had to examine it and tell whether it met his expectations. In case students expressed dissatisfaction, they were asked to explain why they did not like the chosen book and were offered another chance to select among bibliographic records. Again they were offered the physical copy of the book for them to comment on their second choice. After two selections, students were offered all the books described in bibliographic records and were asked to compare their selected books with other available books, commenting why another book would be better or why the book they chose using the bibliographic record was still their preferred one.

Results. Choosing between bibliographic records, the most important elements for high school students were: a note indicating what contents are included in the edition, illustrations, intended audience, and genre. However, when deciding between different copies in a physical form the deciding attributes were: the state of the copy, newer edition and the year of publication, appearance of the book, attractiveness of the cover, the weight of the book, as well as the size and the shape of the letters.

- 68% of students were satisfied with the book that they selected using bibliographic records. When presented with physical copies of all the 11 versions, however, 49% participants would prefer a different book from the one they have chosen.
- Only 33% of participants would use a library catalogue for the purposes of home reading.

Comments.

Research questions: How do high school students select and identify which book is appropriate among 11 different versions of Don Quixote? Which attributes are the most important when they need books for assigned reading? Do they even use the library catalogue?

Adults

Aim. To get a better understanding of which bibliographic elements are important and useful for identification and selection of relevant fiction materials in case of adults.

Data collection technique. Questionnaire and observation

Research Questions. Do a different record design and enriched content have an influence on users’ satisfaction with the chosen book? Which bibliographic elements play a key role when adults select among different versions of the same work using bibliographic records and when they make a choice using physical copies? Based on which elements do adults change their selection when they are given the physical copies of the books?

Study Design. Focusing on fiction, 3 works (each represented with 6 different editions) have been chosen for our test: Quo Vadis (Henryk Sienkiewicz), The Godfather (Mario Puzo), and The Catcher in the rye (J. D. Salinger). For each edition, a physical copy of the book was obtained and three different types of bibliographic records prepared: the first (type A) was copied from the Slovenian union catalogue, while the other two were designed by us and differed in form as well as the set of bibliographic elements. Record type B therefore included some information that was already present in the next generation catalogues as well as some other attributes that users might find interesting such as weight and the colour of the spine, while record type C was based on FRBR.

Procedure. The study was carried out during July and August 2012 with 108 volunteers, who were invited to take part in the study as they were departing from a public library. Each participant in the study would first answer some general questions about the library catalogue and the attributes that were important to him or her when selecting fiction. Afterwards, participants were asked to complete two tasks, first using the bibliographic records and then the actual books.

Based on the six presented bibliographic descriptions for a title, participants selected the one they felt was best for their information need and would hypothetically wish to borrow. After selecting the record, participants were presented with the physical copy of that book and asked to comment on whether they would be happy with their choice. Then the participants would be given the remaining five editions they did not choose with the question whether they would rather select another edition based on a physical copy and why. All the titles as well as the types of bibliographic records were counterbalanced, which means that each participant would
get the three titles in a random order and would be randomly given a different type of bibliographic description for each title, creating all combinations of titles and records.

**Results.** Author, description on the back cover, theme, genre, and the cover presented the key elements in participants’ selection.

- We could also observe that in current bibliographic records, the lack of elements leads users to make assumptions based on the data that is provided (for example, linking the year to the state of the book and the modernity of language, the size of the book with the size of the letters and the density of text, the publishing house with the quality of the translation etc.).

- When participants received the book they selected using bibliographic records, their satisfaction with the book was quite high for all three record types (between 81% and 84%). However, when presented with all 6 physical copies of the book, there were again a number of participants who wanted to change their selection. With traditional bibliographic records, 43% of participants would wish to change the book, while with the enriched records this percentage was reduced: with record type B to 38% and with record type C to 29%. This indicated that even relatively small improvements in bibliographic records (adding an image of a typical page, book cover and identification of contents) could enhance users’ satisfaction by letting them know more in detail what kind of book they can expect, thus closing the gap between the expected and the actual copy.

**Comments.** A relatively small number of people refused to take part in the study and the ones who participated showed willingness to explain their decisions. This may be contributed to the personal approach to each individual.

- Using different bibliographic records in a printed form enabled the participants to be fully focused on bibliographic data without being distracted by other catalogue functions.

- Using video recording or eye tracking could enhance the amount of data gathered in such a study.

**Librarians – focus group**

**Aim.** To establish how well current library catalogues help librarians answer users’ questions and to get the librarian’s perspective on what is important when users choose fiction. Using group interaction we wished to encourage a more in-depth discussion on the topics that would be provoked by sharing of experience.

**Data collection technique.** Focus group

**Research Questions.** Are current library catalogues helpful to reference librarians? What kind of questions do library users pose to reference librarians and what attributes and relationships are most important to them?

**Study Design.** Librarians were asked which questions they could not answer well using the library catalogue and what were the most common user questions and requirements when searching for fiction. Each of the tasks and questions served as a starting point for a discussion.

**Procedure.** Conducted in January 2014, the focus group involved 5 reference librarians from a major public library. Using a combination of tasks and questions as the basis, the focus group took two hours. All the tasks and questions were designed in a way that each participant would first express her view and then the moderator would lead the discussion by presenting more detailed questions and by encouraging the exchange of views.

**Results.** Librarians were generally very satisfied with the current catalogue, but despite their positive and uncritical view on the library catalogue, the conversation revealed that it does not help them answer all users’ questions as it does not include all the needed information and functions. Asking them to list the attributes and relationships that define user needs for certain groups revealed some interesting aspects:

- parents when searching for children books:
  - illustration, typography, reading level
- children: illustrations, page layout
- youth: short description, cover
- high-school students: foreword, full text, abridged edition
- adults: awards, time period
- elders: print size, the weight of the book

**Comments.** Conducting a focus group study, there is always a danger that some individuals will dominate the discussion, thus preventing more quiet participants to express their opinion. Trying to avoid this problem and create a more equal environment, we also designed individual tasks which gave each participant the chance to formulate their answer which were then used as a basis for discussion.

- Focus group presented an excellent base for planning future research.

**Librarians – interviews**

**Aim.** Similarly to focus group, the interviews also aimed at tapping into librarians’ experience with the library catalogue and perceptions of users’ needs they encounter daily. While focus group has its advantages, it may also prevent participants to be completely relaxed and open with their opinions as they might fear what others think of their answers, especially when related to their work. Not really familiar with the method, librarians were also not that keen on participating in a focus group study, but were happy to accept an invitation to an interview about their work.

**Data collection technique.** Interview

**Study Design.** Retaining the same main questions from the focus group, interviews aimed at the same goal but instead of drawing on group dynamics they focused on gaining a deeper understanding of an individual librarian.
Procedure. Six interviews took place in three public libraries during April 2014. The answers were recorded using a tape recorder.

Results. Librarians described the current library catalogue as useful, informative, and convenient. However, the conversation also revealed that when the catalogue is not useful, librarians tend to use various recommendation lists on their webpages or search the web for more information.

- Selecting among different versions of the same work, users will choose the one that looks nicer on the outside, but typically a librarian would present all the available versions to the users, leaving the final choice to them.

- Library users most often search for continuation of a book, parts of a series, or movie adaptations. For different reading levels and purposes, librarians pointed out the following attributes:

  - parents when searching for children books: genre
  - children: illustrations
  - youth: thickness of book
  - high-school students: foreword
  - adults: genre, language, reading level
  - elders: print size

Comments. Compared to the focus group, the answers to our questions were shorter, but participants would also elaborate more on the questions they deemed important. As with the focus group, the interviewer needed to keep a close eye to make sure that the conversation did not drift too far from the main theme.

Discussion

With the longstanding cataloguing practice, it seems that neither librarians nor other users ask (anymore) if the library catalogue gives all the needed information or whether some things are missing. Talking to the participants in our studies it became obvious that librarians as well as users assume that there was something wrong with their search strategy or with their lack of knowledge about bibliographic data. When, for example, one participant chose a different book when she was given all the books in a physical form, she commented: “It all says in the record, it is just that I don’t make out what it means”. Similarly, a librarian in a focus group pointed out that “everything can be found with UDC, it is just a bit complicated”. The conviction that the library catalogue is fine as it is and that the main problem lies in user’s knowledge of the system is a big barrier towards creating a more efficient catalogue. We have made some important steps forward with better display of data, navigation and web 2.0 tools, however, it is the quality and the structure of data that are the prerequisite for a useful catalogue, a catalogue that would be better employed by its end-users.

Although a simple questionnaire would be less time-consuming and easier to analyse, we felt that investigating our research agenda by combining user observation and personal interviews would help us discover information that would otherwise remain hidden. With interviews it was possible to get a better understanding of participant’s choices that in turn gave us the answer to the question which bibliographic elements are important to certain user groups. With our initial research we could see that users’ answers on questions about bibliographic data differed from what we could then observe when users were working with real bibliographic records. That is why our later studies even more carefully and deliberately included various bibliographic records with different bibliographic data for users to work on the chosen tasks.

Two studies, on the other hand, looked at reference librarians as another distinct group of catalogue users. With the intention to encourage a more in-depth reflection on the usefulness and efficiency of a catalogue as librarian’s basic reference tool, a focus group was carried out to engage participants in a discussion. Having some difficulties in recruiting librarians to participate in the focus group, we decided to carry on with the same set of questions and tasks using individual interviews. This way we could also gather opinions and experience from librarians that would otherwise not be able to take part in a focus group due to different factors (distance, nature of their work). While the gathered information from both studies gave a better insight into user’s needs and preferences as viewed through the lens of experienced librarians, we feel that additional studies such as observations at reference desks or tasks similar to the ones we have to other user groups would provide even more information.

Conclusion

Libraries are part of a changing environment and continuous research on what different user groups need is essential if libraries wish to detect and quickly respond to these changes. However, simply asking users what they require or how they select books usually does not give very comprehensive results as people may not consciously recognize the elements they pay attention to; we have seen that even for librarians such questions were difficult to answer as they limited their thinking to the currently available systems and bibliographic data. In case of our studies, observing users as they performed and commented on specific tasks yielded much richer information compared to questionnaire type of answers provided at the beginning of the study.

While our observations were not done on the field, we feel that the tasks were close to a real-life situations (for example, a user at home writes down the books he wishes to borrow, but realises in the library that the desired book or edition is not what he had expected) and therefore reflect some of the issues users are faced with as they use the library catalogue.
Besides author and title there are differences among different user groups in the needed bibliographic elements. We have observed that next to more objective data such as the size of the book or the number of pages, users often selected a specific copy based on more subjective aspects that are not always easy to determine, for example the reading level and the condition of the copy.

In our studies we also observed a substantial gap between the choices made using bibliographic records and those using physical copies. The fact that so many participants would select a different edition if they were choosing among physical copies is a clear indication that more user studies on this topic are needed in order to design more informative bibliographic records. Such research will be needed also for e-book collections where some attributes will become irrelevant (weight, letter size) while other will retain their importance (for example, is there a foreword or a biography included in the book).

REFERENCES


Curriculum Vitae

Katarina Švab is a PhD student and a young researcher at the University of Ljubljana, Faculty of Arts, Department of Library and Information Science and Book Studies.

Dr. Tanja Merčun is a research associate at the University of Ljubljana, Faculty of Arts, Department of Library and Information Science and Book Studies.

Dr. Maja Žumer is Professor of information science at the University of Ljubljana, Faculty of Arts, Department of Library and Information Science and Book Studies.