Interactive Search Profiles as a Design Tool

Maram Barifah(✉) and Monica Landoni

Università della Svizzera Italiana (USI), Faculty of Informatics,
Lugano, Switzerland
{maram.barifah,monica.landoni}@usi.ch

Abstract. *Interactive Information Retrieval* (IIR) research studies how users interact with IR systems and evaluates the users' satisfaction with the retrieval process. Thus, it focuses on how users behave when they interact with IR systems. The involvement of potential users, and access to dynamic and individual information needs are essential elements in IIR studies. User-oriented evaluations investigate the ability of users to engage with a system, thus requiring real users to be involved in the evaluation process. Despite its importance in driving the design of better and more usable IR systems, the user-oriented evaluation approach has been criticised for: (i) being expensive and time consuming in its application, (ii) failing to capture a holistic understanding of the context, as it focuses on users, (iii) delivering not reusable data, often of qualitative nature, and (iv) generating experiments and results that are not reproducible. In this paper, we propose an alternative approach, based on data extracted from *Log File*s (LF) and automatically extracted *Usage Pattern* (UP)s. *Interactive Search Profile*s (ISP)s are then created as tools to guide the design of more usable IIR systems, specifically considering *Digital Libraries* (DL).

Keywords: Evaluation · Digital library · User interaction · Interface

1 Introduction

IIR systems are designed to target users regardless of their numerous and complicated attributes [4]. DLs, as an example of IIR systems, define as "distributed systems with the capability to store various electronic resources and provide convenient access for end users via networks" [33]. Users tend to use DLs to fulfil their information needs. Their interactions are embodied in the information seeking behaviour which are affected by different attributes including expertise in using the system, familiarity, frequency of using the system and knowledge. Lacking expertise and knowledge result in poor usage of the system's functionalities and content identification. Therefore, it is important to enlarge the scope of the evaluation of IIR systems by considering more user attributes [3]. Evaluation of DL assessed the decision makers to determine the usefulness, usability, and economics of a DL [7].

M. Kurosu (Ed.): HCII 2019, LNCS 11566, pp. 18–30, 2019.
https://doi.org/10.1007/978-3-030-22646-6_2

Trischler and Scott emphasised the needs for developing new techniques and tools to redesigning service systems in such a way that users become value co-creators of their own experiences [28]. Such new methods need to be driven by deep understanding of human experience to support the users in co-creating their desired experiences [27]. Collaboration and co-production can lead to open innovation and continuous improvements to increase efficiency, and to develop better interaction experiences [28].

Different approaches can be followed to understand the user interactions and their experiences with the systems. For example, conducting naturalistic studies with real users, and running laboratory-controlled experiments where surrogate users are recruited. Conducting such experiments with test subjects is time consuming, expensive, laborious in terms of organising and running them. Besides, experiences are subjective, and the experiments are difficult to reproduce [1].

Alternatively, researchers may investigate users' digital footprints or UPs recorded in the LF. UPs are a means to represent the search strategies found in the LF [17]. The analysis of the UP enables the researchers to understand the recorded searching experiences [13]. LF analysis can reveal a wealth of information about the user interactions and their preferences. Such analysis does not only results in providing information about the system performance, but also provides a more in-depth analysis of the user interactions with the systems [10].

This work introduces ISP, as a design tool, which can be constructed on top of real data gathered while the users interact with the system. Different ISPs depict the types of interaction a segment of users performs with the system. Such profiles can complement existing user-oriented techniques e.g. personas. The paper starts with reviewing related works, it then describes the research platform, presents the methodology, discusses examples of ISPs, and concludes by suggestion of some practical benefits.

2 Background

Due to the interdisciplinary research domain of the DL covering information retrieval systems, human-computer interaction, and information science, the literature is rich with various evaluation frameworks and techniques [29]. Generally, the existing DL evaluation research can be classified into: general DL evaluation, evaluation of specific components of DLs, and user-oriented evaluation [15,31].

- General DL evaluation studies encompassed the comprehensive evaluation frameworks and models which encourage the researchers and librarians to consider wide range components when evaluating a DL. Scholars e.g. [19,23] suggested to evaluate a DL as a unit by assessing: (i) *content* including digitised resources selection and structure, collection building, repository managing, and preservation. (ii) *policy* counting service, process, quality, access, development, and sustainability. (iii) *technology* considering architecture, algorithms, functionality, interface, and design. (iv) *user* covering perceptions, activities, needs, and tasks.

- Evaluation of specific components of DLs: due to the complexity nature of the DL and its components [7], some researchers contributed to the DL evaluation field by deeply investigating individual aspects. Some evaluation studies [3] targeted the system components and information retrieval performance. Others, considering the metadata [32], and the system interfaces [16].
- User-oriented evaluation: the system performance can be assessed by the established system-oriented approach to measure the effectiveness and efficiency of the underlying search algorithms [26]. The system-oriented approach has been criticised for lacking insights into the user-system interaction and disregarding iterative and exploratory user interactions [20]. Thus, user-oriented evaluation gives the priority to the users by considering their characteristics, information needs and information seeking behaviour [7].
  Consequently, a group of researchers assessed the DLs in relation to the user needs and satisfactions. Such evaluation studies conducted by involving and considering the users and their needs [20]. Three frequent user dimensions considered in the evaluation: usability [30], accessibility [31], and interactions [16].

Accordingly, there is no universal accepted evaluation frameworks or tools, the choice between the available frameworks mainly depends on the evaluators and their aim of the evaluation [7]. Saracevic [22] claims that the impact of such studies are non visible which can be due to the complexity of the DLs, the limited evaluation interest and funds, the lack of the evaluation as a core activity in the operating DL culture [22]. This suggested for more handy and inexpensive evaluation tools.

Evaluation can be conducted at different phases of the DL projects with various purposes [7]. At the initial and during stage of the project, a *formative* evaluation can be conducted to establish goals and to minimise imperfections before release. In contrast, *summative* evaluation is conducted in the final stages to assess the goals are met. *Comparative* evaluation where different systems or components are compared. And *iterative* evaluation aiming to improve the system incrementally [7]. Iterative evaluation is recommended for the DLs as new material are added, metadata are updated and users characteristics changed or new users joined [20,30]. Thus, there is a need for more user-oriented, practical, and inexpensive iterative evaluation tools.

Because DLs are designed to provide support for seeking activities, Sumner [25] emphasised the impact of understanding user, innovative user interfaces and interaction mechanisms on providing better experiences of digital library.

A number of user-oriented design techniques and tools were developed to help designers to make user-informed decisions. Mapping techniques and personas are examples. Such techniques capture the dynamism of processes, and the ongoing interactions during the user experience. Personas, in particular, are a medium for communication which is created or used as an interaction design technique introduced by Cooper [8]. The usefulness and value of personas is increasingly recognised because they offers insight regarding users' attitudes, preferences, and interests [28]. Personas are created fictionally based on extracted attributes from

real users. Courage and Baxter provide idealised attributes list including: identity (demographic data), status of the users to the system e.g. (primary, secondary), goals of using the system, skill set, tasks, relationship, and requirements, expectations, and photograph [9]. Personas assist developers and designers to keep the user present in the design process, and move away from a developer's perspective to a perspective of user [12].

There are different applications of personas: enhancing user testing and evaluation, scenario generation, design exploration, and solution brainstorming. They are usually produced in the initial phases of the design process. Where the real users are complex and non consistent, personas are instead well defined and clear and therefore better suited as a starting point for design work [9]. Personas can be obtained by simple observational methods or in-depth interviews, surveys or studies. Personas are rich but static descriptions of fictive users. Once they are built, the contents of their description are not changeable [12]. The challenge of creating personas is embodied on focusing on a specific segment of the targeted users, or design vaguely for everyone [28]. Persona also has been criticised by being too flat to engage designers, and there is a need for developing characters with richer personalities and better descriptions [12]. Designing for user-experiences is difficult since the experiences do not yet exist. The design situation is therefore only partly known, thus personas need to be invented [28]. Therefore, in this research we make the most of using the existing experiences of the users recorded on the LFs, and presenting ISPs. Such profiles can complement existing techniques for user modelling while offering unique features as described in Table 1, where they are compared with personas.

Table 1. Comparison between persona and ISP

	Persona	ISP
Purpose	representing potential users	representing existing users
Nature	static	dynamic
Data source	qualitative tools e.g. survey	log files
Effort	manually made	automatically made
Cost	costly, time consuming, intensive	cheap, quick, repeatable
Construction	user data collection and analysis	machine learning techniques
Time of use	initial phases of design	after launching systems
Design Effect	reminding about user needs	involving users to improve design

This paper describes how to build and use ISPs, which require a profound understanding of the users searching experiences, the so-called customer journey. The focus is on analysing diverse UPs that occur during the interactions with a system of heterogenous users. We propose to use these ISPs to inform the design

and improvement of interactive systems. The profiles also can be used timely while the users interact with the system where the interactions are captured from the LFs. The proposed methodology will enable to run iterative evaluations, by limiting and possibly avoiding the problem of recruiting participants, and running laborious user studies.

3 RERO Doc

This study is conducted in collaboration with RERO Doc[1]. RERO Doc is a digital library connecting libraries of Western Switzerland as a public service. Figure 1 shows the homepage of the library. The library offers free access to its contents and services. The users can discover the content by simple (1) and advance (2) search functions. Or by navigating the content by the collection (3), institutions (4), content (5), or digitised press (6). The users also may check the latest news of the library (7). The geographical locations of the users are world-wide. Thus, RERO Doc serves a diverse population. LF analysis was conducted aiming to investigate users' interactions and habits recorded in the LFs. Ultimately, the found UPs used to build different ISPs by generating various interaction scenarios.

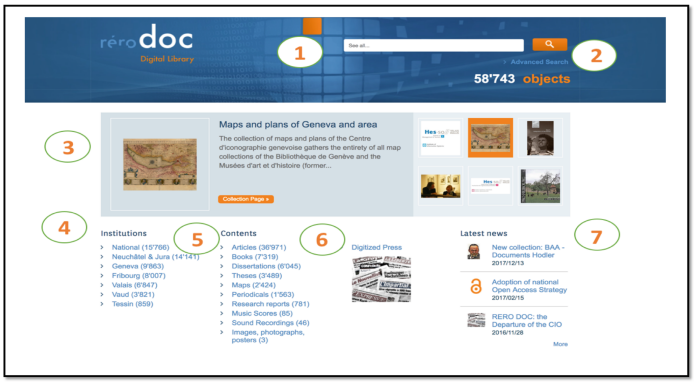

Fig. 1. RERO Doc interface

4 UPs Exploration

For a better understanding of the users journeys on RERO, we draw a hierarchical taxonomy of RERO as shown in Fig. 3. Users may start (B) their sessions by accessing through landing pages including search engines, their emails, or directly through the different functions available on the home page. They can discover (C) the content by searching (G) or navigating (H), After reviewing the search result page, they do have different options of displaying (D) the items. There are different other (E) services the users might implement.

[1] https://doc.rero.ch.

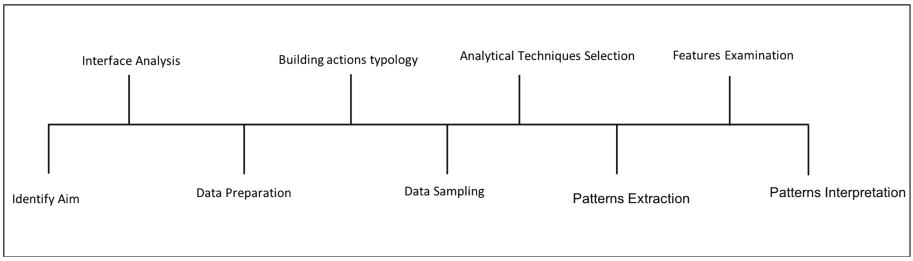

Fig. 2. The framework of exploring UPs in LFs

A dataset of 28 Million records obtained from RERO Doc DL was adopted for our study. Two different unsupervised techniques, namely: K- means, and Birch (Agglomerative) clustering, were applied to data samples of different sizes. The data processing and analysing is described in [2]. Figure 2 shows the framework of exploring UPs.

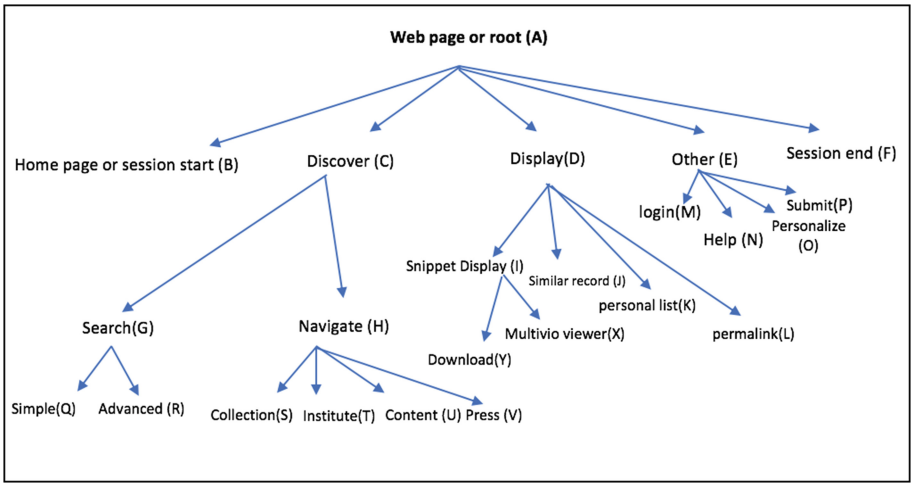

Fig. 3. Hierarchical taxonomy of RERO

This research concluded that UPs can be characterised by: session duration, session starting points, discovering content actions, functions used, and determination session points. Three main UPs were found in the LF, those are: item seekers, navigators, and searchers. Each one of those patterns has sub-patterns. Those patterns and their descriptions are as the following:

1. *Item seeker* (IS): the first category of the UPs is the ISs where they redirected from search engines or emails seeking authorised items. Their sessions characterised by conducting one action download or view items. There are four sub

UPs under this category: Satisfied IS (SIS), Multivio IS (MIS), Average (AIS), and Advance (DIS) item seekers. They are varied in the session durations: short (60), average (60–300), advance (900–1800) seconds. MIS represent the users who prefer viewing items by using Multivio viewer (a viewer application available in RERO Doc) instead of pdf format.

2. *Navigator* (N): this segment of the users discovers RERO Doc by navigating its content which is categorised into: collection, institution, content, or press. Some of the Ns filter and sort the search results by implementing different facets. Their session durations vary between short (60), average (60–300), and long (900–1800) seconds. Their termination actions are: view results list, display item, download item, and add to personal list. Accordingly, we classified them into: Light (LN), Average (AN), Advance (DN), and Press (PN) navigators.
3. *Searcher* (S): the S interacts with RERO Doc by submitting queries through simple or advance search functions. They vary in terms of: session durations, filtering or sorting results, and termination actions. Their session durations vary between short (60), average (60–300), and long (900–2700) seconds. Some of the S filter and sort the search results by implementing different facets. Their termination actions are: view results list, display item, download item, and add to personal list. With some of the Ss, the searching went through many iterations, including reformulation queries. The differences of using the functions could be due to their information needs e.g. KS searchers frequently utilise author and keyword facets. Accordingly, we classified them into: Known item (KS), Simple (SS), Average (AS), Familiar average (FAS), Advance (DS), Familiar advance (FDS), and Sophisticated (PS) searchers.

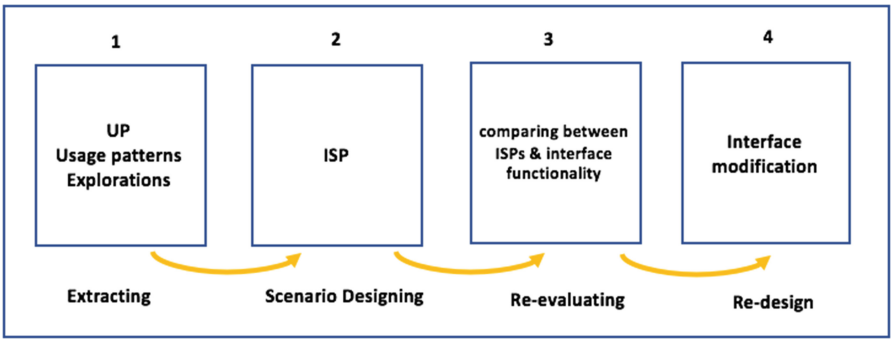

Fig. 4. The implementation of ISPs

5 From UPs to ISPs

Motivated by [5] who emphasised the usefulness of the scenarios for better understanding of the new user situations, we created different scenarios extracted from UPs to be presented in the ISPs. ISPs are specific instances of the UPs, they can represent any of the possible variations inside the same pattern. Figure 4 shows the roadmap of the implementation of the ISP. The process starts with exploring UPs from LF, these can be used to construct different ISPs, and different experience scenarios can be designed accordingly. Finally, such information can assist the designers to re-evaluate their systems.

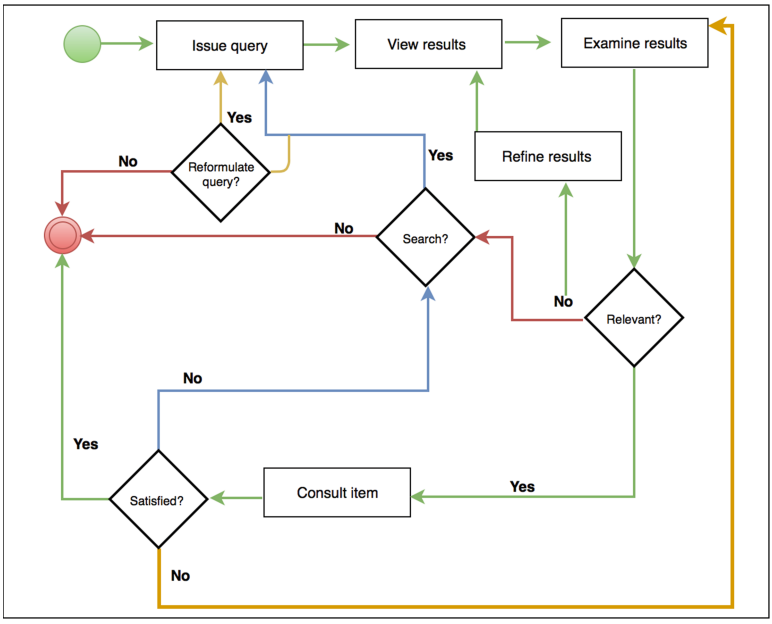

Fig. 5. Searcher model

We developed different interaction models, Fig. 5 is an example of the searcher model. The flowchart illustrates the key actions of the model represented in the rectangles, along with decisions to be taken by the user in diamonds. The visualisation of the interaction models helps us to generate various scenarios.

For example, we identified the UP of Light Navigator, LN, whose session duration last for 60 s and navigate the DL without going beyond the result page. Out of this pattern, we can extract sub LN pattern by simulating all LN features with modification of the search path. For example, instead of finishing up the session with view results page (VRP), a LN may also select a result and apply some functions. Accordingly, slight changes of the termination points also may apply where a LN may download or display the documents. Another example

of deriving different ISPs out of the same UP is that of the KS Searcher. This pattern represents users who usually spend 60 s as a maximum duration and search by author and keyword filters. Their search usually ends with the viewing of the page with the result list. Out of this pattern, a further ISPs can be extracted with different scenarios, what if the KS searchers could not find their information needs? Different paths/tactics might be applied e.g. a KI searcher might return to home page and reformulate the query without filtering the facets i.e. author or keyword.

6 Design Implications

The large-scale logs provided a holistic picture of user behaviour while interacting with RERO Doc DL. Developing different scenarios may help in improving the existent DL interface design. Figure 6 shows the standard result page interface. Consequently, the specific needs of DL users are served effectively and efficiency.

Here we propose some suggestions on how DL may provide mechanism that support user engagement.

6.1 Promoting Different Search Strategies

The usage of the DL depends on the familiarity and experience levels of the users. Stelmaszewska and Blandford [24] investigated how the interface can support users to develop different search strategies. Thus, providing DL interfaces with assistive tools may result in faster, and more successful searching results.

The existing interface design of RERO Doc considers the needs of different searchers by providing simple and advanced search functionalities. In addition to the traditional keyword-based search, DL administrators can enhance the searchers? experiences by the following solutions:

- *Visualising the search results*: Hajra and Tochtermann proposed an approach for visualising search interface by applying external thesauri and suggested terms through machine learning techniques [11]. Consequently, the mental workload of the user is reduces and this results in better search experiences. This suggestion is in line with Cao et al who emphasised the importance of using visual representation rather than textual interfaces to search document collections [6].
- *Providing auto completion function*: although RERO Doc enhances the searching experiences by utilising Auto Correct function, it also could add the Auto Complete or "Popular Terms" function. This might accelerate the searching time [14]. The previous suggested solutions are applicable in the case of the FDS and PS searchers where they reformulated their queries many times during their searching process.
- *Better utilisation of the filter and sort functions*: Niu and Hemminger [18] emphasised the importance of the facets as a supplementary feature for better interaction. They concluded that facets improve the search accuracy for

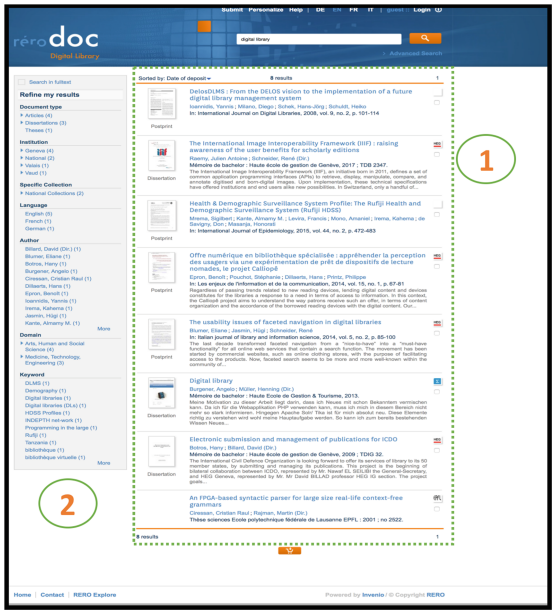

Fig. 6. Result page interface

complex and open-ended tasks by filtering the search results. We found that the KI searchers utilised author and keyword facets more frequently and rarely used the other facets. This indicates that their information needs are more precise than open-ended tasks where different facets where used. Besides, it is worth noticing how not all of the users are aware of the functionality of the facets. As a suggestion, the DL designers might provide a drop-list of the facets instead of having a long static facets bar. The drop-down list might draw the attention of the different available facets.

In terms of sorting the results, RERO Doc offered result sorting function where users can sort their results by: Ascending (SA), Descending (SD) Date (Default), Title (ST), Author(SU), or by including only full-text results. Presenting all the results including abstract, non-full text and full-text documents might causes frustrations for some users and can adversely affect the discovery experience [14]. Although RERO Doc located "search in fulltext" in the upper right corner of the search result page, the sorting functions are rarely used during the interaction processing. Instead, the designers could provide a pop-up window suggesting to limit the search results to include only the full-text results.

6.2 Improving Navigation Experiences

Considering the heterogenous users of RERO Doc, offering navigational icons might improve the navigation interactions. Rahrrovani emphasised the

importance of the icons as a significant feature for better navigation experiences [21]. They concluded that the users with higher mental modules are more compatible with the iconic interfaces. Icons are preferable to abstract text as they are easier to memorise. Thus, providing RERO Doc with navigation icons may assist this segment of users to more effectively interact with the interface.

7 Conclusion

In contrast to the formative and summative evaluation, this paper emphases the importance of the iterative evaluation, particularly when targeting interface design. Iterative evaluation of user interfaces involves steady refinement of the design based on user-oriented evaluation methods. User interactions, as a significant influence factor, need to be considered when refining the interfaces. Based on a wealth and rich user-centred evaluation tool. i.e LFs, this research proposes an iterative evaluation tool. Starting by extracting UPs, we construct ISPs and generate different scenarios to be used for studying design implications and recommend possible improvements. It may be stated that ISP as a data-driven profile, and iterative evaluation tool can be used to continually refine and improve DL interface. The limitations of this approach are embodied in the validation of the found UPs from the LF analysis. Thus, we are planning to conduct a user study to investigate to what extent LFs are informative. We also plan to investigate the usefulness of such profiles for system designers with a future expert study involving experts in DL design and development.

Acknowledgments. The authors thank RERO Doc administrators for their collaboration. Our special thanks to Mr. Johnny Mariéthoz for his effort. And to Mr. Alessio Tutino from BUL (https://www.bul.sbu.usi.ch) for the enlightening conversations on DLs, users and librarians.

References

1. Azzopardi, L., Ìarvelin, K.J., Kamps, J., Smucker, M.D.: Report on the SIGIR 2010 workshop on the simulation of interaction. ACM SIGIR Forum **44**(2), 35–47 (2011)
2. Barifah, M., Landoni, M.: Interactive search profiles as a means of personalisation. In: CHIIR (2019)
3. Behnert, C., Lewandowski, D.: A framework for designing retrieval effectiveness studies of library information systems using human relevance assessments. J. Doc. **73**(3), 509–527 (2017)
4. Belkin, N.: On the evaluation of interactive information retrieval systems. Rutgers University Community Repository document (2010)
5. Blandford, A.: Understanding user's experiences: evaluation of digital libraries. In: DELOS Workshop on Evaluation of Digital Libraries, Padova, Italy (2004)
6. Cao, N., Sun, J., Lin, Y.R., Gotz, D., Liu, S., Qu, H.: FacetAtlas: multifaceted visualization for rich text corpora. IEEE Trans. Vis. Comput. Graph. **16**, 1172–1181 (2010)

7. Chowdhury, S., Landoni, M., Gibb, F.: Usability and impact of digital libraries: a review. Online Inf. Rev. **30**, 656–680 (2006)
8. Cooper, A.: The Inmates Are Running the Asylum. Sams Publishing, Indianapolis (1999)
9. Courage, C., Baxter, K.: Understanding Your Users: A Practical Guide to User Requirements. Methods, Tools, and Techniques. Elsevier, San Francisco (2005)
10. Gooding, P.: Exploring the information behaviour of users of welsh newspapers online through web log analysis. J. Doc. **72**(2), 232–246 (2016)
11. Hajra, A., Tochtermann, K.: Visual search in digital libraries and the usage of external terms. In: 2018 22nd International Conference Information Visualisation (IV) (2018)
12. Johansson, M., Messeter, J.: Present-ing the user: constructing the persona. Digit. Creat. **16**(04), 231–243 (2005)
13. Joo, S.: Investigating user search tactic patterns and system support in using digital libraries (2013)
14. Kay, J.: Improving access to e-resources for users at the university of derby: enhancing discovery systems with library plus 2.0. Insights (2019)
15. Li, Y., Liu, C.: Information resource, interface, and tasks as user interaction components for digital library evaluation. Inf. Process. Manag. **56**, 704–720 (2019)
16. Matusiak, K.K.: User navigation in large-scale distributed digital libraries: the case of the digital public library of America. J. Web Librariansh. **11**, 157–171 (2017)
17. Ndumbaro, F.: Understanding user-system interactions: an analysis of opac users' digital footprints. Inf. Dev. **34**(3), 297–308 (2018)
18. Niu, X., Hemminger, B.: Analyzing the interaction patterns in a faceted search interface. J. Assoc. Inf. Sci. Technol. **66**, 1030–1047 (2015)
19. Noonan, D.: Digital preservation policy framework: a case study. EDUCAUSE Review Online (2014)
20. Petrelli, D.: On the role of user-centred evaluation in the advancement of interactive information retrieval. Inf. Process. Manag. **44**(1), 22–38 (2008)
21. Rahrovani, S., Mirzabeigi, M., Abbaspour, J.: The trained and untrained users' mental models compatibility with the icons of search modules in Iranian digital library applications. Library Hi Tech (2017)
22. Saracevic, T.: Evaluation of digital libraries: an overview. In: Notes of the DELOS WP7 Workshop on the Evaluation of Digital Libraries, Padua, Italy (2004)
23. Shen, R., Goncalves, M.A., Fox, E.A.: Key issues regarding digital libraries: evaluation and integration. Synth. Lect. Inf. Concepts Retr. Serv. **5**, 1–10 (2013)
24. Stelmaszewska, H., Blandford, A.: Patterns of interactions: user behaviour in response to search results. In: Proceedings of the JCDL Workshop on Usability of Digital Libraries Usability of Digital Libraries 2002 (2002)
25. Sumner, T.: Report on the Fifth ACM/IEEE Joint Conference on Digital Libraries – Cyber Infrastructure for Research and Education, 7–11 June 2005, Denver, Colorado. D-Lib Magazine (2005)
26. Tamine-Lechani, L., Boughanem, M., Daoud, M.: Evaluation of contextual information retrieval effectiveness: overview of issues and research. Knowl. Inf. Syst. **24**(1), 1–34 (2010)
27. Teixeira, J., Patrício, L., Nunes, N.J., Nóbrega, L., Fisk, R.P., Constantine, L.: Customer experience modeling: from customer experience to service design. J. Serv. Manag. **23**(3), 362–376 (2012)

28. Trischler, J., Scott, D.R.: Designing public services: the usefulness of three service design methods for identifying user experiences. Public Manag. Rev. **18**(5), 718–739 (2016)
29. Tsakonas, G., Kapidakis, S., Papatheodorou, C.: Evaluation of user interaction in digital libraries. In: Notes of the DELOS WP7 workshop on the evaluation of Digital Libraries, Padua, Italy (2004)
30. Xie, I., Cool, C.: Understanding help seeking within the context of searching digital libraries. J. Am. Soc. Inf. Sci. Technol. **60**(3), 477–494 (2009)
31. Xie, I., Joo, S., Matusiak, K.K.: Multifaceted evaluation criteria of digital libraries in academic settings: similarities and differences from different stakeholders. J. Acad. Librariansh. **44**, 854–863 (2018)
32. Zavalina, O., Vassilieva, E.V.: Understanding the information needs of large-scale digital library users. Libr. Resour. Tech. Serv. **58**(2), 84–99 (2014)
33. Zha, X., Wang, W., Yan, Y., Zhang, J., Zha, D.: Understanding information seeking in digital libraries: antecedents and consequences. Aslib J. Inf. Manag. **67**(6), 715–734 (2015)