"What was this Movie About this Chick?"
A Comparative Study of Relevance Aspects in Book and Movie Discovery

Toine Bogers[1(✉)], Maria Gäde[2], Marijn Koolen[3], Vivien Petras[2], and Mette Skov[4]

[1] Science, Policy and Information Studies, Department of Communication and Psychology, Aalborg University Copenhagen, Copenhagen, Denmark
toine@hum.aau.dk
[2] Berlin School of Library and Information Science, Humboldt-Universität zu Berlin, Berlin, Germany
{maria.gaede,vivien.petras}@ibi.hu-berlin.de
[3] Huygens ING, Royal Netherlands Academy of Arts and Sciences, Amsterdam, The Netherlands
marijn.koolen@huygens.knaw.nl
[4] Department of Communication and Psychology, Aalborg University, Aalborg, Denmark
skov@hum.aau.dk

Abstract. In recent decades, information retrieval research has slowly expanded its focus to address the wealth of complex search requests present in our work and leisure environments. A better understanding of such complex needs could aid in the design of more effective, domain-specific search engines. In this paper we take a first step towards such domain-specific understanding. We present an analysis of a random sample of 1000+ complex book and movie search requests posted in the LibraryThing and Internet Movie Database forums. A coding scheme was developed that captures the 29 different relevance aspects expressed in these requests. We find that while the identified relevance aspects are remarkably similar for complex book and movie requests, their relative occurrence does vary considerably from domain to domain.

Keywords: Query analysis · Book search · Movie search
Information need categorization · Relevance aspects

1 Introduction

The increasing popularity and presence of computers, smartphones and tablets in our daily lives have a correspondingly strong influence on our information seeking behavior. With their use no longer confined to the work environment

G. Chowdhury et al. (Eds.): iConference 2018, LNCS 10766, pp. 323–334, 2018.
https://doi.org/10.1007/978-3-319-78105-1_36

only, searching for leisure has become a major part of the way we use web search engines [19,28]. As a result, the past decade has seen a steady increase in the amount of research dedicated to different aspects of everyday-life information seeking behavior. Savolainen [26], for example, proposed a typology of different types of leisure searchers, while McKenzie [16] introduced a general model of the entire everyday-life information seeking process.

The question of what people search for and what aspects of the desired resources they mention in their requests remains vague in many domains, however, even though this information is crucial for building successful search and discovery systems. For instance, despite the popularity of movie discovery services such as Netflix, our understanding of how people search for and discover, which movies to watch next, is still underdeveloped. Moreover, we lack a good understanding of how relevance aspects across different domains compare to each other. In this paper, we take a first step towards answering these questions and addressing this research gap by comparing two domains: books and movies. We focus on requests that elaborate a searcher's information need more than simple web search queries do, in order to elicit more details about their relevance aspects. Discussion forums are places where we can expect to find people expressing such complex requests. Considering this context, our study is not only related to everyday-life information seeking, but is also part of a growing research community in information retrieval that focuses on complex search tasks [4]. We collected over 120,000 discussion threads from the LibraryThing and IMDB discussion forums and annotated a random sample of 503 book and 538 movie search requests expressed in these threads. We developed a coding scheme for the relevance aspects expressed in these requests and present here an analysis of the results for both domains. Finally, we compare the relevance aspects and the relative distributions of the categories in both domains.

We find that requests in forum posts reveal several relevance aspects related to content and user experience that are rarely considered in system design, but could be addressed using appropriate search functionalities and data sources—e.g., book text, movie subtitles, and user reviews. In this sense, our analysis provides pointers for future system development.

2 Related Work

Query intent, query categorization and relevance aspects. Researchers have previously studied the content of and intent behind queries in order to improve the performance and relevance decisions of search engines. Although relevance is a fundamental concept in information science, very different viewpoints and characteristics evolved over the years. In this study, we use the broader phrase 'relevance aspect' to encompass subject and cognitive relevance expressions in search requests [24]. For web search, query intent analysis has long focused on the three goals (informational, navigational and transactional) first identified by Broder's [6] now classic study. Automatic query classification algorithms are commonly based on these goals [29]. Other studies have broadened the spectrum

of relevance aspects by further specifying the intents [22] or analyzing the content of the query [11]. Studies on academic search engines have adopted Broder's goals to make results comparable [12,15]. They found that searchers generated a lot more informational requests in academic contexts in contrast to web search. This demonstrates that while almost all search environments are now web-based, the domain of inquiry affects the relevance aspects. In everyday-life information seeking, television [8] and music [10,14] are domains where relevance aspects have also been studied.

Relevance aspects in book search. In comparison to other information seeking domains, book search has been studied more extensively. Some studies focus on book selection [18,21,23,27] rather than on book searching. Other studies focus on search strategies in physical bookshops [7] or libraries [17,20] rather than online. Despite the differences, some of the identified relevance aspects are similar to this study.

Our categorization scheme is based on and adapted from relevance aspects in previous studies in the domain of books [13,21,23], but with the aim to cover multiple, related domains. The eight relevance aspects identified by Koolen et al. [13] were elicited from LibraryThing book requests, just as in our study. There is strong overlap with the coding scheme for books by Mikkonen and Vakkari [18], which is based on selections from library catalogs, using four simulated search tasks. Similarly, there is an overlap with the classification scheme reflecting users' multifaceted reading goals developed by Pejtersen [20], based on user-librarian conversations. We base our coding scheme on naturally occurring search requests in forums to avoid the constraints introduced by existing systems and simulated tasks. This resulted in additional aspects compared to Mikkonen and Vakkari [18], as well as different choices in grouping aspects. Our coding scheme also led to a category for search task type.

Relevance aspects in movie search. Movie information needs have not been given the same level of attention as books. In media studies and psychology, the relationship between watchers and movie choices has been studied, for example, the role of age or gender in movie genre selection [3]. For movie selection, Austin [1,2] found that high school students choose movies first based on plot, followed by the actors and then based on friend recommendations. A first categorization of complex requests was performed by Bogers [5], who annotated 400 IMDB forum threads into eight broad relevance categories with 30 sub-level categories. Bogers' study and categorization scheme served as a motivation and background for our analysis of movie requests here.

3 Methodology

In order to conduct a comparative analysis of the aspects that make books and movies relevant to searchers, we collected a representative sample of book and movie search requests. We first describe the data collection process, then the

development of our coding scheme for relevance aspects in the book and movie domains.

3.1 Data Collection

Books. We collected examples of book requests from the online discussion forums on LibraryThing (LT), one of the major social cataloging websites[1]. Currently, there are over 196,000 threads in the LT forums[2], many of which are dedicated to book clubs and reading challenges. The starting point of the data collection process was a 2012 crawl of the first 131,000 LT forum threads posted up to that point. All threads without any empty or hidden first posts were converted to XML, resulting in 115,858 XML threads. As part of the work by Koolen et al. [13], a focused sample of 1,461 threads were manually classified as book search requests, after having been pre-filtered using a simple regular-expression-based classifier, which removed all posts not containing one or more 'trigger' expressions, such as '*suggest*', '*looking for*' and '*which books*'. For the purpose of our analysis, we only extracted the 1,461 first posts in each thread that contain the original requester's book search request[3].

Movies. In order to collect examples of movie requests, the Internet Movie DataBase (IMDB) message boards were chosen. IMDB shut down its message boards on February 20, 2017, so it was not possible to get an exact overview of its size at the time of writing, but in 2015 they contained over 1.19 million threads [5]. Discussion on the IMDB message boards covered a wide variety of topics, but for the work described in this paper, we restricted ourselves to threads where users were most likely to describe their movie-related information needs. Threads that covered movie news, reviews, or discussion of specific movies, actors, director or other aspects were not considered. We restricted ourselves to two message boards in particular: (1) "*I need to know*" (INTK), which typically (but not exclusively) contained known-item requests, where user are interested in re-finding a specific movie already known to them with the purpose of determining the title, using descriptions of the plot or other aspects of the movie; and (2) "*Lists & recommendations*" (L&R), which contained explicit requests for movie recommendations or lists of similar movies with a particular theme.

We combined two different non-overlapping crawls of all threads posted to these two message boards, one originally documented by Bogers [5] and conducted in June 2014, and the second conducted between February 15–17, 2017. The combination of these crawls resulted in 6,320 INTK threads and 634 L&R threads for a total of 6,954 threads. Although the exact proportion of known-item threads in the INTK message board is hard to determine without coding them, the majority consists of known-item threads, which suggests a strong skew

[1] http://www.librarything.com/, last visited December 5, 2017.
[2] According to http://www.librarything.com/zeitgeist, last visited December 5, 2017.
[3] Available at http://toinebogers.com/?page_id=779.

in the distribution of information need types already before coding. Again, all threads without any empty or hidden first posts were converted to XML, resulting in 6,879 XML threads[4]. For the purpose of this paper, we only extracted the 6,879 first posts in each thread that contained the original requester's movie search request. In contrast to the book requests, these movie requests were not pre-annotated as requests or non-requests, resulting in a mix of both types.

3.2 Coding

Development of initial coding schemes. The first phase in analyzing the relevance aspects was the development of initial coding schemes for both domains separately. These initial coding schemes were based on an open coding approach. All five authors acted as annotators and developed their own individual coding on the same development set of book and movie posts. The size of this initial development set had to be large enough to ensure that even infrequent relevance aspects had a decent chance of occurring in that set. Based on the coding frequencies from an earlier book coding scheme [13], we set the size of the development set at 50 posts, as the least frequent coding category occurred once every 27 posts on average. For movies, we set the size of the development set to 75 posts to take into account that two-thirds of all threads from the INTK and L&R messages boards were movie search requests according to Bogers [5]. For each thread in the development and final coding sets, annotators were shown the title and full text of the initial post as well as the group it was posted in.

Calibration of the final coding scheme. The development phase resulted in 10 different initial coding schemes, five for each domain. In total, initial coding resulted in 89 different relevance categories for books and 82 for movies. To arrive at a single coding scheme for each domain, we used card sorting to split, merge, and label the initial categories into a smaller set of unique categories, one for each domain. All grouping or merging decisions where made on the premise to inform information systems that support heterogeneous real-life user requests with different strategies.

After this initial phase, we grouped related categories into top-level aspects. Next, the individual coding schemes were sent around for discussion by all five authors until agreement was reached about the aspects and their labels. The book coding scheme was calibrate *before* the movie coding scheme, which meant that the latter was influenced by our experiences with the former.

In our final discussion round, we attempted to identify similar aspects in both domains and unify the names and descriptions of these aspects so that a unified coding scheme for books and movies could be developed, although some aspects only occur for one domain. Textual descriptions of the different aspects were then added for each aspect along with prototypical examples of each aspect to aid in the final annotation process. None of these aspects are mutually exclusive.

[4] Available at http://toinebogers.com/?page_id=779.

Actual coding process. After calibrating our coding scheme for both domains, each annotator was tasked with annotating 120 book requests and 120 movie requests and provided with a random selection of posts. Posts from the development sets were not re-used for the final annotation. Because not every post was a search request—especially for the movie threads, which had not been pre-filtered—annotators kept annotating until they reached 120 true requests. In practice, this meant between 123–126 book and between 181–218 movie posts were annotated in total. After the first round of final coding, all annotators discussed their experiences. This led to the addition of a Dialogue aspect to the coding scheme as well as clearing up any possible confusion about specific aspects, after which every annotator revisited their 120 requests to harmonize their annotations. Coding agreement is reported in Sect. 4.3.

4 Book and Movie Requests

The final coding scheme (see Fig. 1) includes four top-level categories—Content, Metadata, Context, and Experience—and an other category, which was used to annotate aspects not already covered. The four main categories are further divided into sub-categories. While there is a great deal of overlap, each domain does have unique aspects. For each request, the type of Information Need was categorized as Discovery, Known-item, Sequence & Series or Similarity. None of the categories are mutually exclusive; requests could be assigned to more than one category and information need, although at least one information need had to be assigned to each request.

4.1 Books

In total, 503 unique posts were annotated for the book domain. The majority of requests are long and complex, containing inquiries related to more than one category. For example, the following post includes details about the topic, author background, recency as well as the perspectives used: "*Can someone suggest a book [on Climate change] that's relatively up to date and 'fair to both camps'? ... One, preferably by an actual, reputable scientist, not someone pretending to be a scientist (not making any claims; I can only imagine)? And I would consider myself a fairly dumb layman." (ID 73244).* Requesters mostly asked for Metadata aspects (80.5%), followed by Content (77.9%), mainly including topical and plot information, and Experience (23.1%) or Context (17.1%) aspects. More than half of the threads were discovery-type requests (53.5%), where the searcher is not aware of any books that match the specified relevance criteria. Similarity-type requests were the second-most frequent information need at 39%. Also, requesters often remember when they read a book: "*I read it in the 1960's but it may have been published much earlier." (ID 36142).* Although this information might not be consistent with the publication date, we still decided to classify these requests with the Release date aspect, since it might help provide the system with an approximate time frame. In known-item requests (where the purpose is to re-find

Top-level aspect		Sub-aspect	Description	Books	Movies
What should it be about?	Content	Characters	Books/movies with specific characters or organizations, types of characters or character development	✓	✓
		Design	Books/movies with a particular design, layout, structure or cinematography	✓	✓
		Dialogue	Movies that contain a particular phrase or style of dialog	✓	✓
		Plot	Books/movies with specific plot or narrative elements or scenes	✓	✓
		Setting	Books/movies set a specific location or near distinct geographical landmarks	✓	✓
		Time	Books/movies that are set in a particular time period or around a specific historical event	✓	✓
		Topic	Books/movies that cover one or more specific topics	✓	✓
What kind of properties should it have?	Metadata	Audience	Books/movies that are aimed at a specific audience	✓	✓
		Contributors	Books/movies that have a particular contributor involved (e.g., author, actor, cinematographer, director, illustrator, etc.)	✓	✓
		Format	Books that come in a specific format	✓	
		Filming location(s)	Movies that were recorded in a specific location		✓
		Genre	Books/movies that fall in one or more specific genres	✓	✓
		Language	Books/movies in a particular language	✓	✓
		Properties	Movies with specific characteristics		✓
		Publisher	Books from a particular publisher	✓	
		Release date	The date/period a book/movie was released	✓	✓
		Soundtrack	Movies with a particular soundtrack or sound design.		✓
		Title	Books/movies that have specific (words in the) title	✓	✓
		Version	Specific versions of a book/movie	✓	✓
How will it be used?	Context	Context	Books/movies that describe the context in or purpose for which they will be used	✓	✓
What kind of experience should it provide?	Experience	Comprehensiveness	Books that cover their topic with a certain level of detail	✓	
		Mood	Books/movies that evoke a certain mood, tone or reading/viewing experience	✓	✓
		Novelty	Books/movies that are unusual or quirky, or have novel content	✓	✓
		Impact	Books/movies that have a specific impact on the consumer or that motivate them in a certain way	✓	✓
		Perspective	Books/movies presenting a story from a particular perspective	✓	✓
What type of need is it?	Information need	Discovery	Books/movies that match relevant aspects, where the searcher is not aware of any books/movies that match their search criteria	✓	✓
		Known-item	Describing books/movies already known to the user with the purpose of re-finding them	✓	✓
		Sequence & series	Requests for a set or sequence of related books/movies	✓	✓
		Similarity	Books/movies that mention other items that the requested item should (or should not) be similar to (in some aspect)	✓	✓

Fig. 1. The coding scheme derived for the book and movie domains. The two right-most columns show, which aspects apply to which domain(s).

a book), traditional metadata elements such as title or author are often forgotten, yet design elements are present: "*Can't remember anything about the author or title, and while I can't quite remember I'm pretty sure it was a chapter book, softcover, with a kind of creepy cover illustration." (ID 72482).* Requests like these resulted in adding a Design sub-category for both domains covering these expressions. This relevance aspect is typically not included in traditional book metadata, but could be extracted from cover images by future search systems. We note that our coding scheme covers most of the book appeal elements (pace, storyline, frame and tone) of Saricks [25, Chap. 3], which are based on reference interviews in libraries. We did not encounter explicit mentions of pace, although qualifications such as 'pageturner' can be interpreted as indirectly referring to pace.

4.2 Movies

In total, 538 posts were annotated for the movie domain. Requesters most frequently asked for Content (95.4%) and Metadata (76.6%) aspects. The other categories played only minor roles (Experience 4.1% and Context 1.9%). A clear dominance of Known-item type requests (86.4%) was observed. Often, requesters remembered plot elements up to specific dialogue or sentences: "*I can only remember: (1) Someone ends up mutilated in a bathtub. (1) The villain was an old Italian mafioso complete with hat and suit. One line he said has stuck with me: "What the sh-t is this?" That's all I remember". (ID 225359993).* As the example shows, plot elements are often complemented with a description of characters. Within the Metadata category, 53.2% of all requests reported a Release date, varying from very specific dates—"*I saw an animated Japanese movie in the summer of 1984" (ID 227693731)*—to more vague descriptions, such as "*I saw a movie or TV show at least 15 or 20 years ago" (ID 228781692).*

Some requests expressed movie-specific relevance aspects. For example, posts sometimes mentioned aspects belonging to Properties, such as a particular format (black & white), a specific type of end credits, movie budget, etc. Another example are movie requests that address the Soundtrack or sound design of one or more movies.

4.3 Inter-annotator Agreement

In order to calculate inter-annotator agreement, we arranged for an overlap of 25 posts between successive annotators. Finally, inter-annotator agreement was calculated over a total of 100 overlapping posts. We calculated Fleiss' kappa, because agreement was calculated between different pairs of annotators [9][5]. In the book domain, agreement on whether a post contains a search request is $\kappa = 0.65$, for movies it is $\kappa = 0.83$. The lower agreement for books is probably due to the fact that they were pre-filtered, creating a stronger skew between requests

[5] Agreement scores for all aspects available from http://toinebogers.com/?page_id=779.

and non-requests. With 94 agreed requests and three agreed non-requests, the three disagreements have a large impact on overall agreement. The eight disagreements in the 100 movie posts have a much smaller impact, because the numbers of requests and non-requests are more balanced.

For the relevance aspects, we computed agreement based only on the posts that both annotators labeled as requests. In general, the top-level aspects show substantial agreement ($\kappa > 0.6$) apart from Metadata in the book domain ($\kappa = 0.32$) and Experience and Context (both have $\kappa = -0.03$) in the movie domain. The latter two rarely occur—respectively in 5% and 7% of the requests at least one annotator considers these aspects to be present—so more double annotations are needed to reliably determine agreement.

For the sub-aspects, agreement is substantial ($0.6 \leq \kappa < 0.8$) or strong ($\kappa \geq 0.8$) for several of the Content and Metadata sub-aspects, such as Plot, Dialogue and Publication date or Release date. Agreement on the type of information need is even higher: it ranges from $\kappa = 0.64$ for Sequence & Series to $\kappa = 0.91$ for Known-item in the book domain, and between $\kappa = 0.78$ (Discovery) to $\kappa = 0.82$ (Known-item and Similarity) in the movie domain. Many other sub-aspects in these top-levels have moderate agreement ($\kappa > 0.4$). Experience sub-aspects are rare and have no or slight agreement ($-0.2 \leq \kappa < 0.2$), except Impact and Perspective in the book domain ($\kappa > 0.6$). Across the two domains, 12 sub-aspects have below moderate agreement ($\kappa < 0.4$), most of which occur in less than 5% of requests), 15 have moderate agreement ($0.4 \leq \kappa < 0.6$), 11 have substantial agreement ($0.6 \leq \kappa < 0.8$) and 10 have strong agreement ($\kappa > 0.8$), and 4 sub-aspects do not occur at all. In general, our annotations are reliable for high-level aspects and specific aspects that are very concrete (plot, dialogue, publication date), but reliability drops with increasingly specific and affective aspects.

5 Comparison

Our results indicate that for both domains, relevance aspects are very often Content- or Metadata-related. Book requests include the Metadata aspect more often (80.5%) than movie requests (76.6%), while movie requests mention the Content more often (in 95.4% of all annotated requests) than book requests (77.9%) (see Fig. 2). More significant differences occur for the Experience (in 23% of the book and only 4% of the movie requests) and the Context aspects (in 17.1% of the book and only 1.9% of the movie requests). Such differences are also observable for the sub-categories. For example, twice as many movie requests include Character information (47% vs. 20%) and almost three times as many requests are for Plot elements compared to the book posts (85% vs. 31%). For book requests, Topic is much more common (43%) than for movie requests (8%).

A striking difference that might have influenced our results is the high proportion of Known-item information needs for movies (86.3%) in comparison to Discovery- (53.5%) and Similarity-related (39.0%) book inquiries (see Fig. 3). This was undoubtedly influenced by known-item requests being more common in the

IMDB message boards, although this could very well be a reflection of movie information needs in general.

The next analysis step would be to compare relevance aspects per information need type. For example, comparing Discovery and Known-item book requests, a remarkable difference between the occurrence of Experience aspects (33.5% in comparison to 6.4%) is observable. The same is true for movie requests. For Known-item book and movie requests, the Content aspect (99.4% and 92.7% respectively) is frequently included because of Plot descriptions (81.4% and 92.7%). In contrast, for Discovery book and movie requests, Topic is the most prevalent aspect. Characteristic for Known-item requests in both domains are Publication Date (74.4%) for books and Release Date (58.7%) for movies compared to Discovery requests, where these aspects occur with less than 4%. These preliminary findings of information need characteristics suggest considerable variations in relevance aspects.

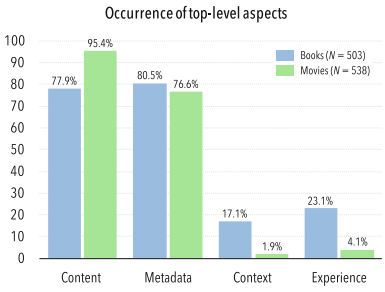

Fig. 2. Occurrence of the top-level aspects in both domains.

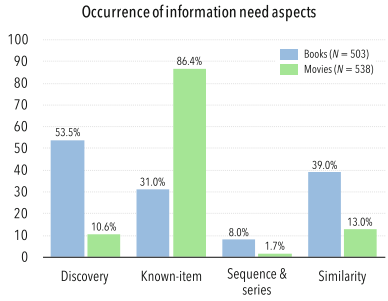

Fig. 3. Occurrence of the information need aspects in both domains.

6 Discussion and Conclusions

Through an analysis of 1041 complex search requests from online discussion forums, we have presented an overview of the different relevance aspects that are prevalent in such requests, but currently not well-supported by search systems. System design for Known-item requests typically considers cases where the user knows important metadata aspects such as title and author, but the forum posts show a different type of Known-item request that require different data sources, such as plot description, movie scripts and book covers to satisfy them. As future work, we would also like to expand our analysis to other domains such as music and game requests. In addition, we would like to evaluate the proposed coding frameworks using content from different forums. This would allow for further validation of our coding framework, especially by focusing on evaluating highly specific sub-categories and affective aspects related to requester's experience where inter-annotator agreement in this study is challenged.

References

1. Austin, B.A.: Motion picture attendance and factors influencing movie selection among high school students. In: Proceedings of 33rd Annual University Film Association Conference (1979)
2. Austin, B.A.: Motivations for movie attendance. In: Proceedings of the 76th Annual Meeting of the Eastern Communication Association, pp. 3–34 (1985)
3. Banerjee, S.C., Greene, K., Krcmar, M., Bagdasarov, Z., Ruginyte, D.: The role of gender and sensation seeking in film choice: exploring mood and arousal. Media Psychol. **20**(3), 97–105 (2008)
4. Belkin, N., Bogers, T., Kamps, J., Kelly, D., Koolen, M., Yilmaz, E.: Second workshop on supporting complex search tasks. In: CHIIR 2017, pp. 433–435 (2017)
5. Bogers, T.: Searching for movies: an exploratory analysis of movie-related information needs. In: Proceedings of iConference 2015. iDEALS (2015)
6. Broder, A.: A taxonomy of web search. ACM SIGIR Forum **36**(2), 3–10 (2002)
7. Buchanan, G., McKay, D.: In the bookshop: examining popular search strategies. In: JCDL 2011, pp. 269–278 (2011)
8. Elsweiler, D., Mandl, S., Kirkegaard Lunn, B.: Understanding casual-leisure information needs: a diary study in the context of television viewing. In: IIiX 2010, pp. 25–34 (2010)
9. Fleiss, J.L.: Measuring nominal scale agreement among many raters. Psychol. Bull. **76**(5), 378–382 (1971)
10. Inskip, C., Butterworth, R., MacFarlane, A.: A study of the information needs of the users of a folk music library and the implications for the design of a digital library system. IPM **44**(2), 647–662 (2008)
11. Jansen, B.J., Spink, A., Bateman, J., Saracevic, T.: Real life information retrieval: a study of user queries on the web. ACM SIGIR Forum **32**(1), 5–17 (1998)
12. Khabsa, M., Wu, Z., Giles, C.L.: Towards better understanding of academic search. In: JCDL 2016, pp. 111–114 (2016)
13. Koolen, M., Bogers, T., van den Bosch, A., Kamps, J.: Looking for books in social media: an analysis of complex search requests. In: Hanbury, A., Kazai, G., Rauber, A., Fuhr, N. (eds.) ECIR 2015. LNCS, vol. 9022, pp. 184–196. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-16354-3_19
14. Lee, J.H.: Analysis of user needs and information features in natural language queries seeking music information. JASIST **61**(5), 1025–1045 (2010)
15. Li, X., Schijvenaars, B.J., de Rijke, M.: Investigating queries and search failures in academic search. IPM **53**(3), 666–683 (2017)
16. McKenzie, P.J.: A model of information practices in accounts of everyday-life information seeking. J. Doc. **59**(1), 19–40 (2003)
17. Mikkonen, A., Vakkari, P.: Readers' search strategies for accessing books in public libraries. In: IIIX 2012, pp. 214–223 (2012)
18. Mikkonen, A., Vakkari, P.: Readers' interest criteria in fiction book search in library catalogs. J. Doc. **72**(4), 696–715 (2016)
19. Pass, G., Chowdhury, A., Torgeson, C.: A picture of search. In: InfoScale 2006 (2006)
20. Pejtersen, A.M.: New model for multimedia interfaces to online public access catalogues. Electron. Libr. **10**(6), 359–366 (1992)
21. Reuter, K.: Assessing aesthetic relevance: children's book selection in a digital library. JASIST **58**(12), 1745–1763 (2007)

22. Rose, D.E., Levinson, D.: Understanding user goals in web search. In: WWW 2004, pp. 13–19 (2004)
23. Ross, C.S.: Finding without seeking: the information encounter in the context of reading for pleasure. IPM **35**(6), 783–799 (1999)
24. Saracevic, T.: Relevance: a review of the literature and a framework for thinking on the notion in information science. Part II: nature and manifestations of relevance. J. Am. Soc. Inf. Sci. Technol. **58**(13), 1915–1933 (2007)
25. Saricks, J.G.: Readers' Advisory Service in the Public Library, 3rd edn. American Library Association, Chicago (2005)
26. Savolainen, R.: Everyday life information seeking: approaching information seeking in the context of 'way of life'. Libr. Inf. Sci. Res. **17**, 259–294 (1995)
27. Serantes, L.C.: "I'm a marvel girl": exploration of the selection practices of comic book readers. In: Proceedings of Annual Conference of the Canadian Association of Information Science (2009)
28. Waller, V.: Not just information: who searches for what on the search engine Google? JASIST **62**(4), 761–775 (2011)
29. Zhou, S., Cheng, K., Men, L.: The survey of large-scale query classification. In: AIP, vol. 1834, no. 1 (2017)