Department of Computer Science
Series of Publications A
Report A-2016-5

Information Search as Adaptive Interaction

Kumaripaba Athukorala

To be presented, with the permission of the Faculty of Science of the University of Helsinki, for public examination in Auditorium XIII, University Main Building, on 12th October 2016, at 12 o'clock noon.

University of Helsinki
Finland

Supervisors
Giulio Jacucci, Unversity of Helsinki, Finland
Antti Oulasvirta, Aalto University, Finland

Pre-examiners
Leif Azzopardi, University of Glasgow, United Kingdom
Max L. Wilson, University of Nottingham, United Kingdom

Opponent
Robert Capra, University of North Carolina at Chapel Hill, United States of America

Custos
Giulio Jacucci, Unversity of Helsinki, Finland

Contact information

Department of Computer Science
P.O. Box 68 (Gustaf Hällströmin katu 2b)
FI-00014 University of Helsinki
Finland

Email address: info@cs.helsinki.fi
URL: http://cs.helsinki.fi/
Telephone: +358 2941 911, fax: +358 9 876 4314

ISSN 1238-8645
ISBN 978-951-51-2516-3 (paperback)
ISBN 978-951-51-2517-0 (PDF)
Computing Reviews (1998) Classification: H.1, H.3.3, H.4, H.5.2, H.5.m
Helsinki 2016
Unigrafia

Information Search as Adaptive Interaction

Kumaripaba Athukorala

Department of Computer Science
P.O. Box 68, FI-00014 University of Helsinki, Finland
kumaripaba.athukorala@cs.helsinki.fi

PhD Thesis, Series of Publications A, Report A-2016-5
Helsinki, October 2016, 122 pages
ISSN 1238-8645
ISBN 978-951-51-2516-3 (paperback)
ISBN 978-951-51-2517-0 (PDF)

Abstract

We use information retrieval (IR) systems to meet a broad range of information needs, from simple ones involving day-to-day decisions to complex and imprecise information needs that cannot be easily formulated as a question. In consideration of these diverse goals, search activities are commonly divided into two broad categories: lookup and exploratory. Lookup searches begin with precise search goals and end soon after reaching of the target, while exploratory searches center on learning or investigation activities with imprecise search goals. Although exploration is a prominent life activity, it is naturally challenging for users because they lack domain knowledge; at the same time, information needs are broad, complex, and subject to constant change. It is also rather difficult for IR systems to offer support for exploratory searches, not least because of the complex information needs and dynamic nature of the user. It is hard also to conceptualize exploration distinctly. In consequence, most of the popular IR systems are targeted at lookup searches only. There is a clear need for better IR systems that support a wide range of search activities.

The primary objective for this thesis is to enable the design of IR systems that support exploratory and lookup searches equally well. I approached this problem by modeling information search as a rational adaptation of interactions, which aids in clear conceptualization of exploratory and lookup searches. In work building on an existing framework for examination of adaptive interaction, it is assumed that three main factors influence how we interact with search systems: the ecological structure of the environment, our cognitive and perceptual limits, and the goal of optimizing the tradeoff between information gain and time cost. This thesis contributes three models developed in research proceeding from this adaptive interaction framework, to 1) predict evolving information needs in exploratory searches, 2) distinguish between exploratory and lookup tasks, and 3) predict the emergence of adaptive search strategies. It concludes with development of an approach that integrates the proposed models for the design of an IR system that provides adaptive support for both exploratory and lookup searches.

The findings confirm the ability to model information search as adaptive interaction. The models developed in the thesis project have been empirically validated through user studies, with an adaptive search system that emphasizes the practical implications of the models for supporting several types of searches. The studies conducted with the adaptive search system further confirm that IR systems could improve information search performance by dynamically adapting to the task type. The thesis contributes an approach that could prove fruitful for future IR systems in efforts to offer more efficient and less challenging search experiences.

Computing Reviews (1998) Categories and Subject Descriptors:

H.1 [Models and Principles]: User/Machine Systems—Human information processing
H.3.3 Information Search and Retrieval
H.4 Information Systems Applications
H.5.2 User Interfaces: User-centered design
H.5.m Information Interfaces and Presentations (e.g. HCI)

General Terms:

Modeling, Search, Experiments, Information

Additional Key Words and Phrases:

Information Retrieval, Information Foraging, Rational Analysis, Adaptive Interaction, Reinforcement Learning, Computational Rationality

Acknowledgements

I have not traveled this exciting yet at times challenging research journey on my own. There are many pillars that supported me and many amazing people who never left my side in this great expedition—they have picked me up when I fell, taught me how to steady my course, and gave me the courage to move forward with confidence. This portion of my dissertation is devoted to expressing my most heartfelt gratitude to all of them.

This work would never have been possible without the financial support of the Helsinki Doctoral Programme in Computer Science (DoCS) – Advanced Computing and Intelligent Systems (Hecse), the MindSee project, the Nokia Foundation, and the infrastructure provided by the Department of Computer Science at the University of Helsinki and Helsinki Institute for Information Technology (HIIT). I am grateful also to the Max Planck Institute (MPI) in Saarbrücken, Germany, and the German Research Center for Artificial Intelligence (DFKI) for hosting me during my research visit. I thank all these organizations for providing the resources to create an ideal environment in which to conduct research.

I extend my warmest thanks to my supervisors, Professor Giulio Jacucci at the University of Helsinki and Professor Antti Oulasvirta at Aalto University. Giulio gave me freedom, while persistently guiding me to progress efficiently. Knowing that he would always be there to back me up with academic and financial resources, I had the courage to explore new horizons of research. Antti made sure that I set my goals high. He guided me to reach high standards while patiently providing me with knowledge as seeds for growth. Together they have tirelessly guided me on the Ph.D. journey.

I have been fortunate to have a great mentor also, Dorota Głowacka, who has supported me throughout. She helped me to explore across discipline boundaries and enhance my research through collaboration. Dorota was also with me at many conferences, giving me confidence.

In addition, I would like to thank the external pre-examiners of this dissertation, Assistant Professor Max L. Wilson and Assistant Professor Leif Azzopardi, for their thorough reviews, which have aided me tremendously in improving this work. Furthermore, I am honored to have Assistant Professor Robert Capra as my opponent.

I am indebted to all the co-authors of the articles that form elements of the dissertation. I am especially grateful to Jilles Vreeken for teaching me a plethora of new skills, not least the secret of keeping calm amid the stress of impending deadlines. Jilles also provided me with insightful feedback on the dissertation despite having a tight schedule himself. In addition, Alan Medlar, Eve Hoggan, Anu Lehtiö, Kalle Ilves, Tuukka Ruotsalo, Ksenia Konyushkova, Samuli Kaipiainen, and Prof. Samuel Kaski have been immensely helpful co-authors.

I would like to thank all my colleagues, past and present, who have provided inspiration as members of the Ubiquitous Interaction group: Imtiaj Ahmed, Oswald Barral, Khalil Klouche, Baris Serim, Yi-Ta Hsieh, Ilkka Kosunen, Matti Nelimarkka, Antti Jylhä, Salvatore Andolina, and Diogo Cabral; within the User Interfaces group: Anna Feit, Daryl Weir, Janin Koch, Mathieu Nancel, and Jussi Jokinen; and at HIIT: Joanna Bergström-Lehtovirta, Luana Micallef, and Antti Salovaara. Special thanks go to Herkko Hietanen for introducing me to the research world and providing me with the wonderful opportunity to work at HIIT, which reshaped my career. Moreover, I thank all of the friendly staff of the University of Helsinki for creating a pleasant work environment. I am grateful as well to my dear friend Laila Daniel, from the Computer Science Department, for constantly motivating me to do my best.

I would never have made this journey without the support of my family. My beloved husband, Dinesh Wijekoon, is the greatest pillar for me, supporting me in tough times, and both my best friend, with whom I have shared the joy of my successes, and the best peer, who read every manuscript that I have written and helped me to improve. I want to thank my loving parents for giving me more than their best. In addition, I am grateful to my two sisters and my in-laws for their immense encouragement and support. Finally, I thank my dearest Lokuge family (Nilmini, Uditha, Samath, and Dinithi) for never allowing me to feel homesick in Finland.

Helsinki, September 20, 2016
Kumaripaba Athukorala

Contents

Chapter 1

Introduction

Search is a fundamental human activity that can take place almost anywhere [107]. Even in the most common of day-to-day chores such as looking for a bus to commute to work and finding basic commodities in a grocery store, we unconsciously engage in search all the time. Among the various contexts in which search takes place, information search in the digital space in particular is gaining more attention with advancements in technology and increased availability of digital content.

Humans are infovores, and our ceaseless hunger for information makes easily accessible digital content an important part of our lives. Today, the digital information space is expanding at an exponential rate [30]. This staggering volume of content poses new problems that require application of special skills and adaptive strategies for foraging large information spaces to find the right information precisely when it is needed [125].

With this dramatic growth of digital information, information retrieval (IR) systems have become an integral part of our lives. Today, we rely on Web-based IR systems to find answers to practically all our information needs. We issue more than 170 billion search queries each month with Web search engines [56] for various needs, ranging from finding basic facts to support everyday decisions, as in checking the weather forecast or determining how to commute to work, to complex knowledge acquisition aimed at gaining expertise over time [107]. Current information retrieval systems are undoubtedly of assistance in satisfying these needs. However, there is evidence that information-seekers struggle when it comes to complex search activities that involve learning or investigations in unfamiliar domains [139, 156]. Such search activities are referred to as exploratory searches [107].

Information search is commonly divided into two broad categories: exploratory and lookup [107]. Exploratory search is a prominent life activity that is often motivated by a complex information problem. This category of searches has been referred to also as general tasks [34], decision tasks [33], subject searches [91], and open-ended tasks [108]. In many cases, exploration begins with an interest in gathering new knowledge of less familiar topics and frequently occurs in an academic context [157, 165]. A typical exploratory task in the academic context would involve a user trying to understand a topic assigned for an essay. Although it often occurs in the academic context, there are many other situations that motivate exploration. A typical example is that of needing to invest in real estate for the first time. We might browse through search-engine result pages (SERPs) to find prices, suitable locations, and other information that could educate us for purposes of making a good decision. A characteristic shared by all of these tasks is that there is a high degree of uncertainty related to what we actually need [147]. Indeed, empirical studies have shown marked differences between exploratory and lookup search behaviors. For example, with exploratory tasks, the user spends more time searching and follows more complex search paths than when performing lookup tasks [32, 104].

When performing lookup searches, on the other hand, users are considered to have precise search goals and a predictable search path. Lookup searches have also been referred to as closed tasks [108], simple tasks [34], information processing tasks [33], and specific tasks [127]. The most distinctive types of lookup involve finding facts (also referred to as factual search tasks) to answer a specific question—for example, the amount of blood the human heart pumps in a minute [14]. For the simplest lookup tasks, the search process can even be automated [32]. Most existing information retrieval systems best support lookups [21].

Of the types of tasks, exploratory searches are considered particularly challenging for the user. One of the reasons for this is that users in this scenario are less familiar with the domain of the search goal [53], which makes it difficult for them to express their information needs and assess the relevance of the search results [160]. Another problem lies in the uncertainty of the search goals and how to reach them [166]. Since the information needs in exploratory searches are broad and complex, users have to browse a broad set of resources; this too is difficult [104]. Furthermore, the user's knowledge and information needs constantly change throughout a search session [150, 157]. All of these factors render exploratory searches challenging.

Furthermore, it has proven rather difficult for IR systems to provide support for exploration [21], not least because of the dynamic nature of the exercise—as noted above, user knowledge, goals, and needs all change throughout the exploration process. Furthermore, distinctly conceptualizing exploratory search poses a great challenge [55,157]. Although numerous models of processes of seeking and retrieving information have been developed [20, 89, 144], they have been largely descriptive and general in nature. For example, the popular berry-picking metaphor presents an analogy between information-seekers who forage patches of information and berry-pickers who travel from one patch to another as they traverse the landscape in search of the best berries from the bushes in each patch [20]. This metaphor is an apt description of the most common information search behaviors. However, it neither provides insights into the reasons behind such behaviors nor can be used to predict which sequences of actions—referred to as search strategies—people will adopt under specific circumstances [16]. Such descriptive models are commonly termed pre-theoretical [81], meaning that they can inform the relationships between factors that support the development of formal and predictive models yet cannot be readily integrated into IR systems for making predictions or explaining observed behaviors.

The focus of the thesis project has been primarily on developing a generalizable conceptualization of information search that answers this fundamental question: *How do people choose which search strategies they adopt in information search?* Proceeding from the conceptualization developed in the thesis project, I build predictive models of search that could be readily integrated into IR systems. The practical implications of the models are then investigated through creation of a prototype IR system that provides adaptive support for various information search tasks.

The work to develop adaptive support for IR is motivated by the problem that most existing support for exploratory information search has been focused on special interfaces, which provide visualizations of keywords [40], categories of results [37], and search time lines [3,50] to help the user understand the structure of the information space. Though specialized interfaces with such visualizations might be useful for exploratory tasks, they are not ideal for lookup tasks. In general, when performing lookup tasks, users are more comfortable with simple interfaces in which search results are presented as a vertical list [73]. In consideration of this, there is motivation to conceptualize exploratory search behaviors and the seeking process through predictive models that enable IR systems to cater to the demands of exploratory and lookup tasks both, simultaneously. The approach I propose in this thesis could address these major challenges facing IR systems and users alike in information search.

My conceptualization is based on *rational analysis* of information search behaviors in lookup and exploratory search activities. "Rational analysis" refers to an empirical approach explaining why the cognitive system is adaptive with respect to its goals and the structure of the environment [39]. I followed this approach because rational analysis has formed the basis for many important models of information search behavior [16, 124]. Furthermore, it serves as an empirical method for predicting how a human cognitive system adapts [5]. In addition, there are computational techniques to implement systematic, rigorous, and general models of rational analysis [63]. I use an existing framework referred to as an adaptive interaction framework (AIF), which adapts rational analysis to the context of human–computer interaction (HCI) [123]. In this framework, human interaction strategies are shaped by three factors: *utility*, what the user finds value in; *ecology*, or the user experience with the task environment; and the *mechanism*, the cognitive and perceptual limits imposed by the information processing system implemented in the human brain [123]. The term "interaction strategy" refers to a set of possible combinations of user interactions with the interface elements. The AIF can be used to recognize all possible strategies the user could perform, which together make up the strategy space. Following the AIF, I examine why we choose one search strategy over another in accordance with variations in circumstances, building my conceptualization of the information search strategies in exploratory and lookup tasks on the basis of the AIF and offering logical explanation as to why exploratory search strategies are dynamic. Furthermore, the framework allows me to implement rational analysis computationally to build a model that predicts the search strategy a rational user would select from within the strategy space. This thesis also contributes two other predictive models for both exploratory and lookup search behaviors on the basis of the AIF. These are designed to 1) predict evolving information needs in exploratory searches and 2) discriminate between exploratory and lookup tasks. Additionally, I offer a suggested approach for integrating these models into an IR system and thereby provide single-interface adaptive support for both categories of search tasks.

1.1 Objectives and Scope

There were three main objectives behind my research: 1) to conceptualize information search—both lookup and exploratory—as adaptive interaction, 2) to build predictive models of information search, and 3) to design an IR system that provides adaptive support for exploratory and lookup tasks. The main outcome of this work can be formulated as five claims.

Claim 1: Exploratory search strategies emerge as an adaptation to ecology, utility, and mechanism under the AIF.

Claim 2: The AIF explains why exploratory search is challenging.

Claim 3: The AIF can be applied to model and explain how users adapt to search results that are either overly broad or narrow with regard to their expectations.

Claim 4: Exploratory searches can be distinguished from lookup searches through user interaction data, and the AIF indicates how.

Claim 5: User performance improves if IR systems retrieve a wide spread of results for exploratory search tasks and a narrower result set for lookup search tasks.

Claim 1: I argue that most problems faced by users in information search are due to current IR systems treating all search objectives in the same manner. This is largely because of the lack of empirical knowledge of how to define search tasks with reference to measurable behaviors. The conceptualization I present, which is based on the AIF, demonstrates the way in which users make rational choices to adapt their search strategies for maximal information gain in a given ecological structure with cognitive and perceptual limits. These adaptive search strategies enable the design of information retrieval systems that can detect when search goals shift from exploratory to lookup and *vice versa* along with when search goals change from being general to very specific.

Claim 2: In this work, I propose that exploratory search is the most challenging type of search activity that users perform today, although it represents a very common purpose for searches. My argument builds on the attributes of exploratory and lookup search strategies as indicated by the AIF. I use the term "search strategies" to refer to the set of sequences of actions, such as issuing a query, browsing the SERP, and opening a document, that users follow when performing search tasks. Rooted in the principle of rationality, the AIF can be used to determine the optimal strategy (or "policy") a user would follow for a given task. By analyzing the characteristics of the optimal strategy, I offer a suggestion as to why exploratory search is challenging for the user. The claim has been confirmed via empirical investigations that involved user interviews followed by observations of users performing natural search tasks. A Web-based survey corroborated the findings from the interviews and user observations. In the thesis, I explicate various factors that motivate the use of electronic search engines and show how users have adapted strategies that differ on the basis of the purpose.

Claim 3: In general, users begin exploratory searches with vague queries using broad search terms. This allows them to obtain cues about new keywords and iteratively formulate queries with more specific terms [150, 157, 167]. Formulating a good initial query is difficult, however, as is reformulating queries when the results are not satisfactory. When users try out queries sporadically, some return results that are overly specific with respect to the knowledge of the user, going into far too much detail. Alternatively, results may be too broad, covering so many sub-topics that it is difficult for the user to get an overview. This is a major challenge, one that leads to users prematurely ending search sessions. In this thesis, I will explain, via the AIF, how user interaction strategies change when the search results returned are found to be overly broad or narrower than the user expected. The primary model I build predicts whether the search results are considered excessively broad or narrow from the way in which the user interacts with them. The model takes the number of search results that the user has seen and the number of clicks on them as input. I have empirically validated this model through a controlled user study and a follow-up free-form study of exploratory search.

Claim 4: Exploratory search strategies are considered to be unpredictable [150, 157]. Since exploration often takes place in unfamiliar domains, we can expect intellectual development with the acquisition of new knowledge [165]. Such developments in user knowledge result in users constantly changing their search goals and strategies in the course of a search session. This dynamic nature of the exercise makes it rather difficult to predict user behavior. Using the AIF, I explain how users adapt their interaction strategies to maximize utility in information search. This explanation allows building of predictive models of the complex exploratory search strategies employed. At the moment, it is hard for IR systems to distinguish between exploratory and lookup search activities. In the thesis, I will identify a set of measurable indicators of common exploratory search strategies, including query length and browsing time, which should assist in IR systems' discrimination between these two fundamental categories of search. The AIF provides a logical explanation for these indicators' ability to reveal exploratory search strategies. Furthermore, I propose a method for detecting these indicators through data logged from user interactions with SERPs. This entails training a classifier to separate between exploratory and lookup search goals while the user is still engaged in the search session. The contribution in response to this claim has valuable implications for the design of adaptive IR systems.

Claim 5: Search engines need to pay special attention to what kinds of results are to be retrieved for the search purpose at hand. Users performing most lookup tasks are satisfied if the results are presented as a ranked list of documents in descending order of apparent relevance with regard to the search query entered. A common approach used by IR systems is to optimize the precision and recall of the search results [26]. However, in exploratory search activities, retrieving the results most closely matching the given search query might leave users trapped in their initial query context [42]. This behavior contributes to users' perceptions of exploratory search as challenging. Through a theoretically oriented analysis, I suggest that user performance of exploratory searches can be improved by adapting information retrieval algorithms to retrieve broader result sets for search queries of a certain nature. I have validated this hypothesis with a controlled study. Considering the models developed in the earlier stages of the research enabled me to propose a suitable approach for designing an adaptive IR system. The proposed approach involves a system that adapts the diversity level of the retrieved results to the predicted task category. If the search task is predicted to be an exploration, the IR system retrieves a broad set of results, covering a wider spectrum of topics. If, on the other hand, the task is a lookup, the system retrieves the documents that best match the search query. In the thesis, I will provide details of the prototype system implemented in line with the approach described above and present results from user studies confirming that the approach indeed helps users to perform both exploratory and lookup search tasks well.

In summary, the thesis builds a conceptualization of information search that allows development of predictive models of search strategies. The models I propose lead naturally towards development of an IR system that provides adaptive support for both exploratory and lookup search tasks.

1.2 Author's Contribution

The main claims made in this thesis are based on seven publications, which are referred to in the text by the Roman numerals used below. The particulars of the publications and my contributions to them are detailed below.

Publication I: Kumaripaba Athukorala, Eve Hoggan, Anu Lehtiö, Tuukka Ruotsalo, and Giulio Jacucci (2013). Information-Seeking Behaviors of Computer Scientists: Challenges for Electronic Literature Search Tools. In: *Proceedings of the Association for Information Science and Technology (ASIS&T)* (pp. 1–11). Wiley. [7]

Contribution: In collaboration with Giulio Jacucci, I identified the need to investigate information search behaviors in the aim of understanding the common practices and open challenges. I designed the study, receiving feedback from Giulio Jacucci and Eve Hoggan. Anu Lehtiö and I conducted the interviews and user observations, and she transcribed the interviews for analysis. I prepared the survey questionnaire and performed both quantitative and qualitative analyses of all data collected, receiving feedback from Eve Hoggan and Tuukka Ruotsalo. I wrote the first version of the manuscript, and all the authors participated in revisions.

Publication II: Kumaripaba Athukorala, Antti Oulasvirta, Dorota Głowacka, Jilles Vreeken, and Giulio Jacucci (2014). Narrow or Broad?: Estimating Subjective Specificity in Exploratory Search. In: *Proceedings of the International Conference on Information and Knowledge Management (CIKM)* (pp. 819–828). ACM. [11]

Contribution: Building a predictive model addressing searches' subjective specificity from observable exploratory search behaviors was my proposal. Receiving feedback from Antti Oulasvirta, Jilles Vreeken, Dorota Głowacka, and Giulio Jacucci, I developed the formal model and designed and conducted the user study. I performed the initial data analysis and validation of the model, while Jilles Vreeken assisted in performing the classification tests. I prepared the first draft of the paper, and all of the authors were involved in the revisions.

Publication III: Kumaripaba Athukorala, Dorota Głowacka, Antti Oulasvirta, Jilles Vreeken, and Giulio Jacucci (2015). Is Exploratory Search Different? A Comparison of Information Search Behavior for Exploratory and Lookup Tasks. *Journal of the Association for Information Science and Technology (JASIST)*, 1–17. Wiley. [10]

Contribution: I formulated the idea of identifying information search behaviors for purposes of distinguishing between exploratory and lookup tasks. I designed and carried out the data collection, receiving feedback from Dorota Głowacka, Antti Oulasvirta, Jilles Vreeken, and Giulio Jacucci. Then, I analyzed the data, while Jilles Vreeken aided in carrying out the classification tests. I wrote the first version of the manuscript, and all the authors participated in the revision stage.

Publication IV: Kumaripaba Athukorala, Alan Medlar, Kalle Ilves, and Dorota Głowacka (2015). Balancing Exploration and Exploitation: Empirical Parameterization of Exploratory Search Systems.

In: *Proceedings of the International Conference on Information and Knowledge Management (CIKM)* (pp. 1703–1706). ACM. [8]

Contribution: Together with Alan Medlar and Dorota Głowacka, I identified the need to calibrate the exploration rate in an exploratory search system, to enable a suitable balance between exploration and moving toward greater specificity (or exploitation). I designed and conducted the data collection, while Alan Medlar ran simulations to identify exploration rates for analysis in the study. Kalle Ilves supported the implementation of the interface, and Alan Medlar designed and developed the back end of the system described in the paper. I analyzed the user satisfaction and performance data, and Alan Medlar analyzed the effect of exploration rate on the number of relevant documents selected. Dorota Głowacka rated the relevance of the user-selected documents. In preparation of the first version of the manuscript, I wrote the "Introduction," "User Study," "User Satisfaction and Performance," and "Discussion and Conclusion" sections, while Alan Medlar and Dorota Głowacka drafted the "System Overview" and "Modelling Document Selection" sections. All authors participated in revisions.

Publication V: Kumaripaba Athukorala, Alan Medlar, Antti Oulasvirta, Giulio Jacucci, and Dorota Głowacka (2016). Beyond Relevance: Adapting Exploration / Exploitation in Information Retrieval. In: *Proceedings of the International Conference on Intelligent User Interfaces (IUI)* (pp. 359–369). ACM. [9]

Contribution: I proposed an approach to adapting exploration versus exploitation in IR systems on the basis of the search task type. I trained a classifier to predict the search goal and integrated this into a search engine that was developed by Alan Medlar and Dorota Głowacka. Then, I designed and conducted the user study, receiving feedback from Alan Medlar, Antti Oulasvirta, Giulio Jacucci, and Dorota Głowacka. Alan Medlar and I together analyzed the data and wrote the "Results" section of the manuscript, and I drafted the first version of the other parts of the manuscript. All of the authors contributed to the revisions.

Publication VI: Tuukka Ruotsalo, Kumaripaba Athukorala, Dorota Głowacka, Ksenia Konyushkova, Antti Oulasvirta, Samuli Kaipiainen, Samuel Kaski, and Giulio Jacucci (2013). Supporting Exploratory Search Tasks with Interactive User Modeling. In: *Proceedings of the*

Association for Information Science and Technology (ASIS&T) 50(1), 1–10. Wiley. [133]

Contribution: I, in cooperation with Tuukka Ruotsalo, proposed an interface for interactive intent modeling and an approach to its integration into search systems. I implemented system logging to collect user interaction data, while Tuukka Ruotsalo, Ksenia Konyushkova, and Samuli Kaipiainen implemented the retrieval algorithm, intent model, and interactive visualization, respectively. I designed and conducted the user study, receiving feedback from Tuukka Ruotsalo, Dorota Głowacka, Antti Oulasvirta, Samuel Kaski, and Giulio Jacucci. Finally, Tuukka Ruotsalo and I analyzed the data and prepared the first version of the manuscript, with all authors participating in revisions.

Publication VII: Dorota Głowacka, Tuukka Ruotsalo, Ksenia Konyushkova, Kumaripaba Athukorala, Samuel Kaski, and Giulio Jacucci (2013). Directing Exploratory Search: Reinforcement Learning from User Interactions with Keywords. In: *Proceedings of the International Conference on Intelligent User Interfaces (IUI)* (pp. 117–128). ACM. [64]

Contribution: I was involved in the design of the interactive visualization and the system features described in the paper. I designed the user study, receiving feedback on this from Dorota Głowacka, Tuukka Ruotsalo, Samuel Kaski, and Giulio Jacucci. Ksenia Konyushkova, Dorota Głowacka, Tuukka Ruotsalo, and I conducted the user studies. I analyzed the user responses from the questionnaire and wrote the first version of the "User Experiment" section. All authors took part in revisions to the rest of the paper, which Dorota Głowacka had the primary role in the writing.

1.3 Structure of the Thesis

Elaborating upon the picture of the two types of search tasks—exploratory and lookup—that I have begun to paint in this introductory chapter, Chapter 2 provides a thorough analysis of existing definitions and theoretical views on search tasks, their differences, and state-of-the-art techniques designed to support information search. Chapter 3 provides a detailed description of the scope of this thesis and of the objectives for it. Then, Chapter 4 presents my conceptualization of information search built in accordance with the AIF. The models that were developed in this research are presented in Chapter 5. Chapter 6, in turn, presents the approach developed to provide real-time support for both exploratory and lookup search. In the final chapter, I revisit the principal claims addressed by the thesis and further analyze the validity of the main research findings. The thesis concludes with a discussion of possibilities for future work.

Chapter 2

Background

Information search is an important activity that is most often performed over the Web with Web-based IR systems. Design of systems to support efficient retrieval of information that satisfies the user's specific information needs has been investigated in several disciplines, among them IR, cognitive science, human–computer interaction, and machine learning (ML). Many investigations began with classification of search tasks on the basis of several factors that contribute to information needs of various kinds. Most commonly used classification systems resulting from this work use exploratory and lookup categories.

In this chapter, I review previous research on exploratory and lookup search tasks. Section 2.1 provides an overview of existing studies that characterize the user behaviors and information needs displayed in various search tasks. The discussion in Section 2.2 addresses factors that make exploratory search more challenging than other search tasks. Section 2.3 presents contributions—with origins in multiple disciplines—aimed at improving user performance in exploration. Other theoretical models of search tasks are discussed in Section 2.4. A summary of previous work and the challenges remaining concludes the chapter, in Section 2.5.

2.1 Current Understanding of Information Search

Figure 2.1 presents a well-known classification of search tasks, which was first proposed in Marchionini's seminal paper on exploratory search [107]. Although the term "exploratory search" entered widespread use after the publication of Marchionini's paper, exploration is not a new phenomenon by any stretch of the imagination. Information search tasks have commonly been grouped in various ways under the general categories of exploratory

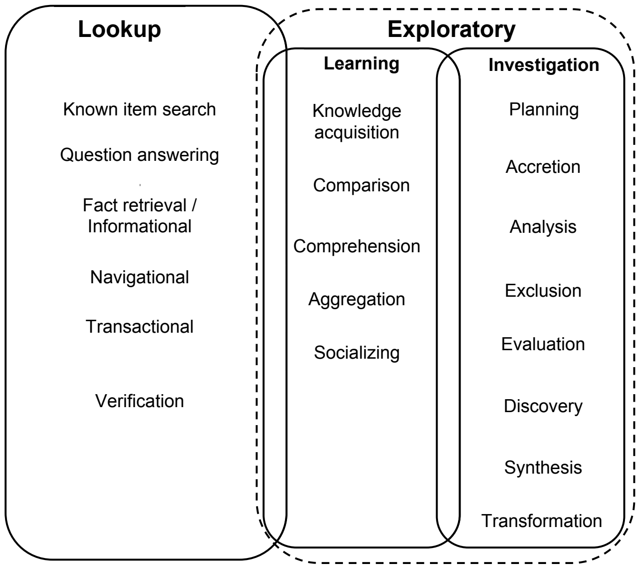

Figure 2.1: Categorization of search activities falling under exploratory and lookup tasks (based on Marchionini's work [107]). Overlap between categories' bounding boxes indicates interaction between those task types.

and lookup tasks. Since lookup searches can be embedded in exploratory searches and *vice versa*, it is difficult to delineate a clear boundary between the two categories. The interplay visible between search tasks is indicated by the overlapping bounding boxes in Figure 2.1. To facilitate ready understanding, I categorize the attributes of exploratory and lookup search tasks as discussed in prior studies into three groups: attributes arising from 1) the *task description*, 2) the *search process*, and 3) *user perception*. Table 2.1 provides a list of the attributes, arranged in terms of these three groups. This section frames the definitions previously used for exploratory and lookup tasks with respect to the attributes listed in the table.

2.1.1 Exploratory Search

Several experimental studies have been conducted, using various attributes in their definitions of exploratory search [163]. The usual general definition refers to the investigation and exploration of information spaces for the purpose of learning and making discoveries [128, 149, 165]. The attributes categorized in Table 2.1 serve as a high-level conceptualization clarifying our understanding of the diverse definitions of exploratory search.

The attributes of the *task description* are the characteristics associated with what motivated the exploration. Common attributes in this category are the *goal*, search *topic*, and degree of *uncertainty*.

Goal: In exploratory tasks, the search goals are open-ended or general in nature, and they are poorly defined [51, 52, 96]. Here, the term "general" refers to conceptually broad goals with no specific target. They are commonly associated with learning or gaining new knowledge and/or investigation [89, 107, 149, 157]. Often, explorers in an academic context are motivated to achieve a higher level of intellectual development within the search domain [149, 157]. Another common goal in exploration is comparison among several topics/areas [96, 107]. An example expressing a comparison goal is that of a student who is considering a university for graduate studies, exploring and comparing all possible universities in a certain geographical region [4]. For most tasks with this attribute, there is no specific answer that satisfies the information need; rather, there are many suitable results, which vary in their degree of relevance [53, 89, 109, 157]. For reason of these fairly loose constraints [136], users in this scenario target multiple documents [149].

Search topic: In many situations, topics for exploratory search arise through discussion/interaction with other people. For example, a professor may ask a student to learn about an emerging research area [11], or someone who has learned of a family member having been diagnosed with a disease may be interested in investigating that disease in more depth [169]. In both scenarios, another person has motivated the search. In exploratory search, the topics are general and multifaceted, which means that they cover many concepts [92] and most often involve a broader domain, with several subtopics [53, 54, 128]. Many prior studies have examined exploratory search tasks that require finding information on an unfamiliar or at least less familiar topic [76, 118, 168].

Uncertainty level: When performing exploratory tasks, the information-seeker is uncertain about what queries to make, what results are relevant, and where to begin the search [107, 157]. Some researchers have referred to this phenomenon as the user facing difficultly in determining in advance which information is required for addressing the need [89]. Previous work shows that during the initial iterations of exploration, the user tries to make sense of the search domain to reduce the degree of uncertainty [53, 157].

The attributes of the *search process* are the characteristics of the information search behavior while the user is engaged in search. Three attributes of the search process are considered here: *duration*, search *path*, and *collaboration*.

Duration: Exploration is considered a longitudinal process [157]. According to the literature, it might involve multiple query iterations and multiple search sessions [128, 157], continued over a long span of time [54].

Search path: With exploratory tasks, we cannot identify a single and direct path that leads to the desired results [157]. As the user keeps exploring, knowledge and the information need continue to change [35, 54, 153]. There are also changes in the searcher's motivation and interests [96, 97]. These dynamic factors render it impossible to predict at the outset what kind of queries the user might issue, what links will be followed, and when the user will terminate the search. The path in exploratory search indicates a browsing-based strategy; that is, the user navigates through broader areas of information rather than focusing on a single, narrow information patch.

Collaboration: Many exploratory searches are prompted by a discussion or other interaction with someone else, so it is unsurprising that the information-seeker might interact with several people in the course of the search process [115]. There could also be many people interested in the outcome / the findings. The following example is an exploratory task posed in a prior study that involved interaction with an external person:

> Your great granny's doctor has told her that getting more exercise will increase her fitness and help her avoid injuries. Your great granny does not use the Internet and has asked you to create an exercise program for her. She is 90-years old. Put together two thirty-minute low-impact exercise programs that she could alternate between during the week. [[169], p. 4]

When performing this task, the searcher might interact with another person, who may have more knowledge of the topic. Exploration could involve collaborating in embarkation on the journey, during the exploration itself, and in the presentation or finalization stages of the search process [94].

User perception refers to how the user subjectively assesses his or her performance of the task. *Subjective complexity* is an important attribute of user perception [99]. In general, users perceive exploratory search tasks to be difficult. In some guidelines, perceived complexity is not considered a key attribute, however [96]. Some researchers have proposed that exploratory search tasks are perceived to be complex because of the lack of support provided by existing information retrieval systems [154], but, at the same time, there are several works that articulate exploratory search tasks as complex problems [118].

Table 2.1: Attributes of exploratory and lookup search tasks.

Attributes	Exploratory search	Lookup search
	Task description	
Goal	Learning and/or investigation, comparison, open-ended task, abstractness, poor definition, multiple-item target	Answering a specific question, clearly defined criterion, navigation to a known target, precise result set
Search topic	Broad or general topics, multifaceted task, less familiar or unfamiliar topics, assignment/motivation from another person	Known-item search, already known specific topic, narrow area
Degree of uncertainty	Uncertainty about the search queries made and of the results' relevance and where to find the information	Certainty about what kind of information to expect, carefully specified queries, precise results, minimal need to examine the results
	Search process	
Duration	Long duration, continuing over many query iterations and potentially several search sessions	Shorter duration, few iterations, termination immediately upon finding of the answer
Search path	No predictable or structured path, dynamic path, combination of browsing and focused search but leaning more towards browsing	Predictable search path, possibility of automated search process, returning of discrete and well-structured objects
Collaboration	Engagement with other people during the search	Search mostly by the individual user
	User perception	
Subjective complexity	Tasks that are considered not very easy	Variable perceived complexity: easy or difficult

2.1.2 Lookup Search

Lookup is a more basic kind of search, returning discrete and well-structured objects, such as specific Web sites or definitions [157]. The most common lookup tasks are targeted at facts and answering a specific question [14]. Marchionini has listed six search activities under the lookup category (see Figure 2.1): fact retrieval, known-item search, navigation, transaction, verification, and question-answering. In some domains, such as library science, the concept of known-item searches is used to refer to lookup tasks in general. Table 2.1 provides a list of common attributes of lookup tasks in comparison to exploratory tasks.

Goal: In lookup tasks, the search goal is very precise, with a specific question and a clear target item in the user's mind. The literature refers to these tasks also as focused searching [157], because the information-seeker is focused on finding a specific target that he or she may just think exists [79]—for example, a particular recipe, a Web site for downloading specific software, or the viewer rating for a certain movie [101].

Search topic: Lookup searches are focused on one topic. The information-seeker may or may not be familiar with that topic, but it is typically very narrow in either case. An example would be finding the answer to the question "what is the most common dog breed in Finland as measured by the number of registrations?" [14]. Most of the time, lookup search topics involve making an everyday decision [79].

Degree of uncertainty: Because the search goal in lookup searches is very precise, the information-seeker is certain about what kind of information to expect. Lookup tasks are sometimes referred to as structured tasks, because the user can clearly express what type of information is useful, concepts and relations in a query, and criteria for relevant documents [147]. A user performing such a task can easily identify the relevant documents with minimal effort when examining the results [107].

Duration: Lookup searches tend not to last as long and generally involve either one or two query iterations [86, 107]. A typical lookup search begins with the user expressing the information need as precisely as possible in order to reach the right neighborhood of information. This is followed by fast browsing and following the link for a few results that look relevant, then settling for the most appropriate item [69]. Prior studies have shown that a lookup task continuing for a longer time and having multiple query iterations is an indication of struggling [14, 72].

Search path: Most common lookup tasks involve simple search paths. Also, it is possible to automate the search process in cases of simple lookup tasks [32]. However, there are broader lookup tasks, in which, while the search goal is precise and the user can determine easily whether he or she has found the answer, the search process is more complex and may involve several paths. One example is finding information on various antivirus software applications and their prices [14]. Some scholars considering lookup tasks that involve thinking or understanding rather than simply locating an item have referred to these as interpretive tasks [89]. Lookup tasks of this nature are more focused and goal-oriented than exploratory tasks, yet they may involve locating several results before an answer is arrived at.

Collaboration: The motivation for most lookup tasks is internal to the user [79]. It is unlikely that another person is involved in the search process, since the search task is more straightforward. Therefore, lookup tasks generally are performed by one user alone, with no collaborators involved.

Subjective complexity: Lookup tasks of the most basic type are perceived to be easy [13], but there are lookup questions that cannot be answered directly. Such tasks are viewed as complex [14]. In consequence, the information-seeker's opinion on the task's complexity is not a good attribute for use in separating between exploratory and lookup tasks [13].

2.2 What Makes Exploratory Search Difficult

Exploratory search has been found to be naturally challenging for users, and it is difficult also for IR systems to offer support for exploration. There are several common reasons for this: the information-seeker lacks the knowledge necessary for formulating search queries that clearly express the information need [20]; exploratory search is a highly dynamic and longitudinal undertaking wherein the user knowledge, search goals, and information needs may change throughout the search process [11]; and there is no proper working definition of exploratory search [157, 161]. These factors are discussed in this section of the chapter.

Lack of knowledge: People who engage in exploratory searches are generally unfamiliar with the search domain and unaware of key terms that might express their information need [157]. Most IR systems present a ranked list of documents in descending order of their relevance to the search query made, with the aim of optimizing the precision and recall of the search results [107]. Retrieving the results best matching the search query could trap the user in the initial query context, and this may, in turn, contribute to the user's perception that exploratory search is challenging [42].

Dynamic knowledge, goals, and information needs: Studies have shown that the knowledge possessed by the information-seeker has a significant impact on search strategies [90, 116, 134]. In exploratory search, user knowledge is subject to constant change. For example, at the beginning of a search session, the information-seeker might have a very vague idea about the search topic [89], but after examining some useful documents, he or she may have a better understanding of the search topic and related terminology [157]. Along similar lines, the process might deviate from the initial search topic to a different topic or to a lookup search targeting a specific result. Users should be able to expect the type of support from the IR system to take such changes in knowledge and information needs into consideration [11]. Detecting dynamic changes of this kind and adapting accordingly is a great challenge in IR system design [21]. Additionally, the dynamic nature of the exploratory search process is a key reason for it being seen as challenging.

Lack of definition: As has been mentioned in Section 2.1, there are various definitions of exploratory search, because all the scholars investigating this topic have introduced their own definitions [165]. Although there is overlap between some of the existing definitions, it is difficult to determine a concise set of of agreed-upon properties to conceptualize exploratory search. The multifaceted nature of exploration contributes further to this problem [158]. The difficulty in identifying when a search task has actually become exploratory is another element complicating conceptualization of exploratory search. Some searchers may begin with certainty about what they wish to find, but the search process might expose them to an unfamiliar area that triggers exploratory search behaviors [161]. Another concern with the current definitions is that they do not refer to quantitative behaviors that can be empirically analyzed [16]. Although many information search models exist, they are mostly descriptive or qualitative [20,94]. They are of use for understanding what kind of user behavior is to be expected in search tasks of various sorts; however, more quantitative models of measurable search behaviors are needed to inform the design of IR systems [21]. Existing definitions cannot be directly applied in IR systems to this end, for recognition of exploratory or lookup search activities [21]. This is one of the main challenges in designing systems that support diverse information search goals.

2.3 Support for Exploratory Search

In this section, I review the contribution of various research communities to improving the support for exploratory search. The goal behind these contributions has been to provide features that address one or more of the challenges discussed in Section 2.2. Table 2.2 presents features with poten-

Table 2.2: Features or approaches to address the challenges encountered in exploratory search.

Challenges	Features/approaches addressing the challenges
Lack of knowledge	Query suggestions, result categorization, visualizations of the information space, facilitation of collaboration
Dynamic nature (of knowledge, goals, and information needs)	Adaptive systems, visualizations, task-management support
Lack of definition	Studies of search behaviors

tial to address the challenges in exploratory search. This review examines various algorithms, visualizations, search systems, and behavioral studies.

2.3.1 Support to Gain Knowledge

At the beginning of an exploratory search, the information-seeker usually lacks domain knowledge. Therefore, IR systems should assist the user in gaining knowledge more rapidly by aiding in formulation and swift refinement of search queries, providing summaries of results through result categorizations or facet-based presentation, and using other visualizations. When performing search tasks of this type, information-seekers often consult someone with better domain knowledge when they are having trouble. Providing support for direct collaboration through the IR system holds potential to improve user performance. Existing systems that provide such support are discussed below.

Query suggestions: All the commonly used IR systems, including Google, Yahoo, and Bing, allow the user to express the information need via natural-language statements or in the form of keywords [157]. In the initial stages of exploratory search, user-defined queries tend to be vague and imprecise in consequence of the lack of user knowledge [107]. It has been found that only 25% of search queries made during exploratory searches are successful [139]. Query suggestions help the user navigate to possibly more relevant or interesting areas of the information space and learn about key terms [64]. One of the first approaches proposed by the IR community involved query expansion/suggestions based on relevance feedback [135]: users mark documents as relevant or non-relevant, and the system then develops a query model and updates it on the basis of the features present in the marked documents. Empirical studies showed that users benefit from such techniques [88]. However, evidence from later user studies showed that users rarely provide relevance feedback, because the cognitive load of selecting relevant documents is high in comparison to typing a new query [88]. In response, scholars investigated implicit means of obtaining relevance feedback [69, 83]. It has been found, though, that query suggestions derived from this relevance feedback often lead to context traps [88]. In another approach, called query by example, the user submits examples of relevant documents, after which the system suggests related queries or documents [84]. However, this technique has been shown to be more suitable for narrowing the scope of a search query than for exploring diverse aspects of a given topic [157]. Nowadays, many Web search engines offer query completion support through menus appearing below the query input box. Such real-time query formulation support techniques have been found to be effective when the user is uncertain about how to express the information need, yet this technique does run the risk of query drift [162].

Result categorization: Result categorization is the organization of search results into meaningful groups [37, 49]. This grouping helps the user to make sense of the information space [74]. There are two popular result categorization methods: clustering and faceted categorization.

In clustering, results are assigned to groups in accordance with their similarity. Typically, similarities between the documents are measured by considering common words and phrases [49]. Clustering algorithms can automatically identify newly emerging topics or information categories in any text collection, which helps the user to understand the themes in the search domain [74]. While offering these advantages, sometimes clustering may result in unpredictable and poorly labeled or unordered categories that could confuse the user. Studies have revealed that users prefer understandable hierarchies with categories that manifest uniform levels of granularity [126].

In faceted categorization, the search results are organized into meaningful categories that reflect the relevant concepts in the search domain [170]. Faceted classification systems allow the user to explore collections of documents by applying multiple filters [146]. Such systems classify information elements along explicit global dimensions, enabling the classifications to be accessed and ordered in several ways. Facets are usually created manually, but the documents are automatically linked to the individual facets. Since the number of global features is often large, the former process can rapidly become overly demanding because users have to go through numerous options [170]. Improved modeling of the user behavior and needs would allow reduction in the number of facets and thereby enhance user experience.

Visualizations: The lack of success exhibited by relevance feedback techniques is often attributed to user interface design failing to provide feedback in a convenient manner at suitable levels of granularity. In response, diverse systems have been designed to support user feedback, including intelligent user interfaces that assist the user in comprehending an information space [18, 40], visualizations that summarize results for purposes of faster relevance judgment [85, 95, 110], and interactive visualizations that allow the user to indicate the direction for exploration [64]. These systems are very useful but designed only for exploratory tasks. For lookup tasks, users find such visualizations to be a distraction, and they prefer simple interfaces in general [73]. Switching between search engines dedicated to specific task types could be expected to entail extensive cognitive overload [124]. Therefore, complex visualizations are not always useful for all search tasks, and a single system that supports all tasks is preferable.

Collaboration support: Many forms of collaboration can be seen in exploratory searches [113]. In the initiation stage of exploration, an external individual might be involved in setting the goals for the search [94]. There can be several collaborators, making various contributions to the search process [67]. By allowing the information-seeker to pool cognitive resources of several types through collaboration, IR systems could improve user performance in exploratory search conditions [115].

Collaboration support can be provided in various ways—for instance, by allowing collaborative generation of ideas [6], query formulation [114], search result exploration [103, 140], and sharing or bookmarking of "favorite" content [46, 87, 112]. Also, there are interfaces that allow users to interact with each other. For example, the social web allows a user to see what other users are viewing at the moment and helps users communicate with each other [66]. The SearchTogether plug-in is another solution that enables remote users to collaborate, either synchronously or asynchronously, while searching [114]. All these applications have proven effective for exploratory searches.

2.3.2 Adaptive Support for Dynamic Changes

One of the grand challenges for IR systems is prediction of the actual information needs of the user [21]. Given that exploratory search is a highly dynamic process, it is beneficial for systems to be able to predict when information needs may evolve and to provide adaptive support, while also helping the user to understand how the search process has been developing.

Adaptive systems: Text classification has been a popular topic in machine-learning research for decades. Applications dealing with the problem of online adaptive learning have appeared only relatively recently [62]. Most examples of text-stream applications involve email classification [36], detection of email spam [102], and sentiment classification [24]. Various adaptive learning strategies have been employed in this domain, with some of the individual methods used being case-based reasoning [28] and ensembles, either evolving or with explicit detection of changes by means of change detectors [1, 25, 68]. However, most of these contributions have no direct implications for the design of adaptive search systems. Adapting to the gradual change in either user interests or data distribution, such methods cannot predict when the actual information needs of the user change as search progresses. Recently, there have been attempts to predict changes in a searcher's topic knowledge at different stages of search from behavioral variables [105, 172, 173]. The predictive power of these models in real-world IR settings is questionable, though, because they have not been tested in actual IR systems.

Visualizations and task-management support: Exploratory searches typically involve multiple search sessions and information gathered from several sources [157]. On many occasions, users backtrack in other stages in their search sessions [107]. Therefore, users benefit from seeing how their search topics, their interests, the queries, and other factors related to the search goal evolve or otherwise change over time [3]. Furthermore, users require tools that enable them to revisit previously encountered items with ease [157].

Some visualizations present information in time lines that enable the discovery and exploration of patterns. For example, Lifelines2 [155] presents patients' medical records and test results along a time line. There are systems that visualize images related to events via browsing on a time line, thereby allowing the user to build a narrative [3]. Such visualizations provide an overview of activities that have taken place in the longer term and aid in making sense of a broad field of information that the information-seeker has been exploring in the course of a longitudinal search process.

There are systems that help the user manage the information. For example, Hunter Gatherer provides an interface with which Web users can collect information from various Web pages, represent it as a collection, and edit that collection [174]. In the same category are systems that construct maps of search concepts that help searchers to see the relationships among various concepts [164].There are also visualizations that support the comparison of parallel streams of search results retrieved by means of multiple search queries [93]. Such visualizations have been verified as useful for making sense of the information space.

2.3.3 Studies of Search Behaviors

There have been various studies aimed at understanding how users adopt strategies and behaviors that are specific to their search *goal*. Various factors may influence search strategies: *search goal*, task *difficulty*, *complexity*, and user *knowledge*. Prior studies of these factors have contributed to identifying quantifiable characteristics of exploratory search. They thereby offer potential for improvement to IR systems through building of empirical models to identify search tasks. In this section I review prior studies of search behaviors.

The *search goal* is the primary reason for a user's interaction with an information search system [89]. In numerous studies, researchers have manipulated the preciseness of the search goal definition and investigated how it affects user behavior. An early study of encyclopedia use by novices introduced two types of tasks [108]: "closed tasks," with precise search goals, and "open-ended tasks," with fuzzy search goals and no definite boundary. The results indicated that with open tasks, novices have difficulty in formulating search queries, take longer, and perform more query reformulations. In another study, scholars investigated the navigation style of novice and expert Web users with known-item search and subject search goals [90], where "subject search" is similar to open tasks. The results indicate that the number of nodes visited, the number of keyword searches performed, and the frequency of clicking on various buttons are influenced by the search goal. Similar studies qualitatively analyzed the information-seeking strategies of Web users with three search goals, termed *factual*, to do with finding a definitive answer in response to a precise search goal; *interpretive*, or configuring an answer for a less precise search goal; and *exploratory*, which involves broadening of knowledge with open-ended search goals [89]. The results suggest that users performing exploratory tasks spend considerable time reading a page returned in the search results in order to determine its relevance. These studies indicate that users behave differently when the search goal is not as precise. We can conclude from these findings that the various terms used—"open-ended tasks," "subject search," and "decision tasks"—refer to the exploratory category of search activities [53].

In other studies, Web search goals have been categorized as informational, navigational, and transactional. Researchers investigated how the *navigational* and *informational* nature of search goals affect cognitive styles [116, 121]. In some studies, external evaluators manually classified search queries collected from search-engine logs by these three types of search goals and investigated how to distinguish among the goal types on the basis of query properties [31, 79, 132]. This work has provided useful findings. However, the log data were assessed by external evaluators, and their evaluation may not fully reflect the intent of the user; hence, the evaluations are rather unreliable [132].

Difficulty and *complexity* are two other important factors that influence user behavior. Task difficulty is always considered a subjective element that depends on user perceptions [99]. Task complexity, in turn, is measured with both objective and subjective approaches. It is difficult to distinguish between subjective task complexity and task difficulty, because they are both assessed by the performer of the task in line with familiarity and degree of uncertainty with respect to the task requirements [23,32,148]. In contrast, objective task complexity is more readily distinguishable from difficulty. It is commonly measured in terms of the number of sub-paths involved in the search process [32]. Tasks with a single determinable path that could be easily automated are commonly referred to as simple tasks, while tasks wherein the process and information requirements are indeterminable are typically categorized as complex tasks. The literature suggests that exploratory search tasks may have high objective task complexity [157].

Several authors have categorized tasks on the basis of the search goal and their complexity or difficulty. For example, Web search tasks are categorized in consideration of the preciseness of the search goal, objective complexity, the product (is the outcome factual or instead intellectual?), and level (whether the document is judged to be a whole or a segment) [104]. Although this categorization does not take the characteristics of exploratory and lookup tasks into account, it is intuitive and shows that there are tasks with mixed characteristics—such as those involving specific search goals but high complexity. Research along similar lines analyzed how task difficulty and two types of search goals—open and closed—influence search behavior [106]. The findings suggested that closed tasks and difficult tasks are associated with long dwell time, a metric for the time expended in reading the documents retrieved. In other work, researchers explored how task difficulty can be determined from information search behaviors by assigning easy and difficult closed informational tasks [13]. They found that as tasks become more difficult, users tend to make numerous search queries, visit results in large numbers, and spend more time on search result pages. Similar studies demonstrate that users engaged in exploratory search tasks display corresponding behavior [72, 107, 159].

The *knowledge* possessed by the user is another factor that influences the information search behavior [99]. Prior studies reveal that Web experts rely heavily on query-formatting tools, while domain experts with less experience of Internet use are heavily reliant instead on terminology and avoid query formatting [77]. Several studies have been carried out in pursuit of greater understanding of how cognitive strategies are influenced by the level of domain knowledge, expertise with the Web, and task type [82, 116, 134]. These studies have yielded qualitative evidence supporting the claim that Web experts follow cognitive strategies that differ clearly from those of novices when exploratory search tasks are involved.

In summary, previous studies point to various information search behaviors, related to task completion time, number of queries, dwell time, and number of resources followed, among other factors, that are affected by task type. However, they have not considered two aspects that are important with respect to the design of IR systems. Firstly, they have focused on Web search rather than IR system use. Hence, many measurements employed, such as the number of unique search engines used, are less informative. Furthermore, there are marked differences between Web search and IR system use, because IR systems constitute a special environment with a specific dataset [80]. Secondly, most of these studies examined search behaviors at the level of the entire search session, rather than that of the first query iteration. If one is to adapt IR systems to different task types, it is important to have measurements of search behaviors that allow the IR system to predict the task type as early as possible.

2.4 Theoretical Views on Search

There are several distinct theoretical approaches to modeling of exploratory search. This section presents three relevant theories: information foraging theory, the berry-picking theory, and the utility maximization theory.

2.4.1 Information Foraging Theory

Information foraging theory (IFT) explains the exploratory search behavior of humans [124] in a manner that borrows from optimal foraging theory—a theory in biology that predicts how organisms obtain energy from their environment—and rational analysis [38]. Rational analysis is an empirical method developed to explain why a cognitive system is adaptive [5]. According to rational analysis, humans optimize their behaviors to maximize reaching of their goals. Through IFT, we are more able to understand, predict, and improve humans' interaction with information. The theory encompasses several quantitative models of user search, with the key idea being that decisions on what to do are made on the basis of the expected information gain. In the process of searching and learning more about the content, the user is continuously updating the "information scent"—i.e., his or her estimate of the information to be gained by selecting a particular item. Information scent, in turn, affects the choice of whether to investigate an element or not. The theory makes predictions as to how information gain, expressed as a function of time, changes with the interface design [48]. When search results are unordered, information gain is a linear function of time. When they are ordered, it shifts to a diminishing returns curve. Scholars have used IFT to explain how the presentation technique (for instance, result clustering) changes information gain rates and when is the optimal time to stop searching. More recently, researchers have applied the concept of information scent to predict the rankings of links on various Web pages [60].

2.4.2 The Berry-Picking Metaphor

The berry-picking metaphor proposes that information-seeking behavior is analogous to picking berries in a forest, where berries are scattered about, on various bushes, and must be picked singly [20]. This is similar to an information-seeker gathering fragments of information from an information space. When moving through the information space, the information-seeker obtains cues that aid in the navigation. The berry-picking theory emphasizes the dynamic needs in search rather than the act of searching itself.

According to the berry-picking theory, when information-seekers encounter new information, they gain new ideas and directions to follow. Proceeding from this new information, the searcher formulates new queries. As the search progresses, the desired outcome and user perceptions of relevance are subject to change. Such a dynamic search process is described as an "evolving search." The information need cannot be satisfied by a single final retrieved set. Rather, the information-seeker is involved in a series of actions of browsing, gathering information, learning of new terminology, and query reformulation actions. This process may continue until an end point of redundancy is reached. The strategy explained in the berry-picking theory is the most commonly employed strategy in exploratory search. The core action in exploratory search is understanding or making sense of fragments of information [157]. In the berry-picking metaphor, the information seen by the searcher influences subsequent actions. In exploratory search, the information encountered influences the knowledge, adding to it and leading to significant changes in the search strategies.

2.4.3 Utility Maximization

The utility maximization theory, which originated in economics, explains why users prefer one particular set of items over others [16]. In economics, this is a problem consumers face [152]: how should I spend my money in order to maximize my gain? Information retrieval systems have exploited this theory in many ways—for instance, for determining the way in which to rank search results [61,131] and for predicting user behavior [16,17]. The fundamental principle behind the associated models is computation of the costs of an IR system and estimation of the gain or benefit for the user [151]. On the basis of these costs and benefits, the system determines the most profitable user behavior or expected search strategy / system features [17]. Very early models predicted the ideal balance of the amount of time a user should spend searching and how much time the system should spend on searching [45]. The probability ranking principle (PRP) is another useful formulation based on utility maximization. It is used to determine the order for search results by considering the costs and benefits of ranking one search result above another [131]. All of these models are useful yet not directly focused on exploratory search.

2.5 Summary and Open Challenges

I will now draw together the review of information search research that 1) conceptualizes exploratory and lookup search tasks, 2) highlights the factors that make exploratory search more challenging than lookup, 3) presents techniques to provide additional support for exploration, and 4) proposes theoretical models of search. The following concluding remarks can be made on the basis of this review:

- Exploratory and lookup search tasks can be conceptualized by considering attributes of the task description, the search process, and user perception.

- Three factors render exploratory search challenging for both the information-seeker and the search system: 1) the information-seeker lacks domain knowledge; 2) the knowledge, search goals, and information needs are dynamic; and 3) there is no proper working definition.

- Today's IR systems use several techniques to support faster acquisition of domain knowledge: query suggestions, result categorization, visualizations, and provision of collaboration support.

- Adaptive systems, visualizations, and task-management support are some of the techniques already in use to address the dynamic nature of exploration.

- Various studies have been carried out to inform understanding of search strategies, in attempts to propose a definition of exploration.

- The berry-picking metaphor, information foraging theory, and the utility maximization theory all provide theoretically oriented views on search that could explain some of the search strategies unique to exploration.

Although all of the works discussed above have made valuable contributions to improving user performance in information search, there are several open challenges with respect to supporting information search.

Those striving to develop systems or techniques that support knowledge gain face various problems. Designed to help the user gain knowledge, result categorization and visualization techniques represent a departure from the familiar list-based search interface. While special interfaces of this nature might be useful for exploratory tasks, they may not be ideal for lookup tasks. In addition, research shows that, in general, users prefer the simple interfaces used to support lookup tasks [73]. Therefore, for the best of both worlds, users may have to switch between systems for exploratory and lookup tasks. However, there is a large amount of interplay between exploratory and lookup searches, which renders it difficult for the information-seeker even to ascertain what kind of search task is being performed and whether switching to a different system is necessary. Thus an important open challenge becomes evident: *how to design IR systems that work well for both exploratory and lookup tasks.*

Systems that propose adaptive support for performing search tasks often either need a long training period before they can detect the gradual change in user behavior or require users to state explicitly that they are planning to conduct an exploratory search. This points to a second challenge: *how to predict the information need and search goals while the user is still engaged in the search task.*

There are many models to predict dynamic attributes of exploration such as the information-seeker's knowledge, perceived task difficulty, and the complexity of the task. However, it is difficult to interpret the performance of these models in connection with real search tasks, because they have not been integrated into actual IR systems. Another concern is that many models are largely conceptual in nature: they do not make predictions or explain observed search behaviors; rather, they are descriptive. Such descriptive models cannot be readily integrated into IR systems. Hence, another challenge remains: *how to build predictive models and integrate them into actual IR systems.*

Although there are several useful theoretical perspectives on information search behaviors, such as information foraging and the utility maximization theory, they consider only a few aspects of why people apply certain strategies. Information search strategies are influenced by many factors, among them the user's existing domain knowledge, experience with the IR system, and the distribution of the results in the information space. Previously developed theoretical models do not empirically explain how these factors influence search strategies. This leads to our final challenge: *how one can theoretically and empirically model all the factors that shape search strategies.*

The goal for the thesis project was to address these challenges and thereby deliver a more satisfactory search experience.

Chapter 3

Research Questions and Method

Information search may be initiated for any of a wide variety of purposes. Among these, exploratory search is rapidly gaining importance as knowledge becomes available in greater and greater quantities via the Web and knowledge bases. Greater understanding surrounding both exploratory search tasks and how to support them better is necessary for improving user performance and satisfaction.

In this chapter, I start by introducing the research questions for the thesis. Then, I provide an overview of the research strategy applied in view of these questions and describe the methods used in the studies conducted. The emphasis here is on the application of a research method that gradually led towards designing an IR system that provides better support for exploratory and lookup searches.

3.1 Research Questions

In the introduction (Chapter 1), I provided an overview of the main claims addressed in the thesis before stepping back to highlight the open challenges connected with exploratory search, in the background chapter. The claims described emerged from a series of studies that were motivated by these open challenges and conducted to address each of the research questions presented in this section.

In pursuit of the objective of guiding the design of information retrieval systems that could support both exploratory and lookup tasks, equally well, I approached the problem in terms of three themes: 1) conceptualizing and understanding information search, 2) modeling and predicting information search behaviors, and 3) providing real-time adaptive support for exploratory and lookup search tasks. These themes were derived from the challenges in exploratory search that I discussed in Chapter 2. The research questions (RQ1–RQ5) are framed in relation to these three themes.

RQ1: *How can information search strategies be conceptualized as a rational adaptation?*

Lack of a proper working definition is one of the key obstacles to the design of IR systems intended to support exploration. Most of the existing definitions are descriptive in nature, rather than encompassing quantifiable or measurable elements. Although a few theories, such as information foraging and the utility maximization theory, propose some quantifiable factors by means of which search can be conceptualized, many factors that could influence the information search behavior are not covered by these theories (as discussed in Chapter 2). We need a systematic conceptualization that frames all the factors that influence search behaviors to explain why and how exploratory search strategies differ from lookup strategies. The work to answer the first research questions responds to this problem by using the AIF, which is an extension of rational analysis (see Chapter 4), to conceptualize information search strategies.

RQ2: *Do information-seekers still need more support for exploration, even with the existing tools and techniques?*

Over the last decade, many techniques, tools, and systems have been proposed to support exploration. This leads one to wonder whether the information-seeker already has sufficient support for performing challenging search tasks more efficiently. The open challenges that I identified through the analysis of prior work shed some light on this question (see Section 2.5). As that section illustrates, it is important to examine the usefulness of the latest tools from the user's perspective. I set out to examine various purposes for which users initiate search so as to identify the purposes that still present a challenge. This should answer the question of whether exploratory search is a prominent purpose and whether users still find it to be challenging.

RQ1 and RQ2 fall under the theme of conceptualizing and understanding information search. By answering these two questions, I identified the following three questions, which are oriented towards designing adaptive information search systems.

RQ3: *How can one predict the dynamic changes in the subjective specificity of information needs in exploratory search?*

As discussed in the previous chapter's presentation of background, one of the key challenges in efforts to support exploration is the dynamic nature of the endeavor: both the knowledge and the interests of the user are subject to constant change. Moreover, whether the results of a query are informative is highly subjective. What is informative to one user could be too specific or less relevant to another. Generally, users initiate exploration with vague queries using broad search terms, an approach that allows them to obtain clues for subsequent reformulation of the queries, with specific terms [157]. When users try out queries exploratorily, some queries return results that are overly specific with regard to the knowledge of the user, going into far too much detail, while other queries may return excessively broad results that cover many sub-topics. Over the course of a search session, the user might also gain knowledge of the search domain, whereupon the information needs and interests may deviate toward a new area. Subjective specificity refers to specificity of search results with respect to the user's actual information need. This subjective specificity of the results differs both between users and internally to an individual user, depending on intent and accumulated knowledge about the domain. If we were able to model subjective specificity from implicit user interaction data, the results could inform the prediction of dynamic changes in the search process. When considering RQ3, I investigated the possibility of modeling such subjective changes in information search in response to simple observable behaviors, such as result clicks.

RQ4: *Can we separate exploratory search from lookup in the course of searching?*

Although there exists a research corpus of ample size on understanding information search behaviors, most of the major IR systems today do not provide adaptive support for such tasks. One reason is lack of empirical knowledge of how to distinguish between exploratory and lookup search behaviors in IR systems. The work on the fourth research question addressed this issue via systematic and rigorous analysis of measurable information search behaviors, referred to as behavioral indicators. The behavioral indicators include readily observable user interactions with the search-engine result pages. They encompass query length, query duration, scrolling depth, and many more indicators. My objective in response to RQ4 was to propose an approach to differentiate between exploratory and lookup search activities while the user is still engaged in the search session. An answer to this question enables deriving implications for the design of adaptive IR systems.

RQ3 and RQ4 are categorized under the theme of modeling and predicting information search behaviors.

RQ5: *How can adaptive support for exploratory and lookup tasks be provided?*

Answering the final research question involved developing a working prototype of an adaptive search system that supports both exploratory and lookup searches. This addresses the observation that most IR systems target only lookup search and present a ranked list of documents in descending order of the relevance the system judges them to have to the search query issued, with the aim of optimizing the precision and recall of the search results. While most of the support thus far proposed for exploration is focused on special interfaces or visualizations, users prefer the simple interfaces typically used to support lookup tasks. The work in response to RQ5 addressed this issue through development of an IR system that dynamically adapts its parameters to the search goal. The model proposed in response to RQ4 was applied to separate between exploratory searches and lookup searches.

RQ5 falls under the theme of providing real-time adaptive support for exploratory and lookup search tasks.

Although the research questions are focused on information search in general, the research strategy involved implementation primarily in the academic information search setting. Hence, the claims have been validated only for academic information search. However, we discuss how the findings can be extended beyond academic search.

3.2 Research Strategy and Methods

The thesis has been written to examine information search in exploratory and lookup search activities, so as to support the design of IR systems that improve user performance. The aim for this work as a whole is to build a grounded understanding of information search through modeling of search behaviors. The models were designed to serve the purpose of developing an adaptive search system. The set of studies conducted along these lines indeed led ultimately to this goal, with development of a concrete system. This section of the chapter provides an overview of the research strategy.

Experimental research methods allow the researchers to design tasks by varying conditions in order to investigate a hypothesis [57]. There is a wide spectrum of experimental research methods available in the field of human–computer interaction, ranging from controlled laboratory-based user studies to more recently developed unmoderated online assessments. It is important to analyze the pros and cons of these methods systematically, for identification of the method best suited to validating the hypotheses chosen in the thesis project. Applying the primary criterion in selection of a research method involves maximizing three features of the measurements [111]: 1) their generalizability, by improving the validity of the results across the population of users; 2) precision, by controlling for the extraneous factors (factors that are not being studied); and 3) realism, by making the situation or context within which the study is conducted resemble the context where the relevant actions naturally occur as closely as possible. As these features all can interfere with one another, it is important to identify the features that are most desirable in light of the hypothesis at hand. To this end, my research involved studies that each had their own focus.

Because the objective in the first stage of the research was to gain a thorough understanding of the relevant information search behaviors, the first studies were focused on maximizing the realism. Accordingly, I initiated this line of research with a set of qualitative case studies for illuminating the natural information search behaviors. Then, to improve generalizability, I conducted quantitative surveys. In the second stage, in which I developed user models, more controlled empirical studies were needed, to identify the features for the models. At that stage, I focused on increasing the precision by minimizing the effect of external factors. I conducted controlled laboratory studies for this purpose. Later, to validate the models and the system proposed, I designed less controlled, more free-form empirical studies. My objective with these was to validate the realism of the models and proposed system. This group of studies provided insights into search behaviors, supported the designing and validation of user models, and aided in investigating the proposed system in a realistic setting. At this juncture, I will offer an overview of these methods. Chapters 4, 5, and 6 provide more details about each of the three stages.

Mixed-methods case studies were utilized in the initial investigation of information search behaviors, which answered RQ2. The data sources were interviews, diary-type logs, and observation of natural search sessions. The use of interviews was a direct approach to get evidence as to how information search behaviors have evolved over time with the availability of digital content. The logs and user observations maximized the realism of the situation or context within which search is performed [111]. Though this approach restricted the investigation to a smaller sample, it provided rich information about the naturalistic information search behaviors.

A *Web-based survey* was used as a follow-up method to improve the generalizability of the findings from the case-study work [111]. It helped to compensate for the small number of participants in the cases examined and increaseed the diversity of the group.

Controlled laboratory studies were developed for the later experiments, to control external factors that could affect information search behaviors. Discussed in Chapter 2, there are many confounding factors that can lead to search strategies similar to exploratory search strategies being manifested, among them task difficulty, complexity, and user knowledge. To validate the models built to predict exploration, one must generate search tasks similar to exploratory and lookup searches while controlling these external factors. To this end, I designed laboratory studies that motivated the participants to perform both exploratory tasks for learning or investigation activities and lookup tasks that involved finding targeted documents. The exploratory tasks designed for these studies were, in general, centered on learning about a less familiar topic for a given time and then writing an abstract about what was learned about the assigned topic. Several of the exploratory tasks also involved answering a set of questions created by domain experts. Such tasks allowed me to situate the participants in an exploratory search context even though the motivation for the search was not natural.

Self-motivated free-form studies with less control had to be created to enable observing of more naturalistic exploratory search behaviors. For this purpose, free-form explorations were conducted on topics that the participants were actually interested in learning about. The approach I followed was similar to the four steps in Borlund et al.'s [27] guidelines for simulating work tasks to evaluate interactive information retrieval systems: formulation of task description based on a personal need or simulated need, relevance assessment of documents by users and a panel of external reviewers, reformulation by the user with reference to the first findings, and reformulation with reference to the third findings. This approach helps to improve the scientific rigor of the search tasks. Analogously, for investigation of lookup behaviors, a range of tasks was created on the basis of the categorization proposed by Marchionini [107] (see Figure 2.1). Self-motivated free-form studies of this nature improve the realism of the search context [111]. A panel of external experts assessed the relevance of the user retrieved documents to the topic of the search task. At the same time users also provided their subjective feedback on the findings. Users had the freedom to reformulate information need (using search queries) as many times as they require. Since all the studies were conducted in the laboratory, I was able to improve the precision by controlling for extraneous factors that were not being studied, such as the influence of social context and of email or other messages appearing that could distract a searcher.

In all of the studies, academics predominated in the sample population. The main reason for this is that exploratory tasks often focus on learning or investigation activities, which are natural and seen more commonly within an academic context [165]. All the studies also included gathering of subjective feedback on performance and satisfaction, and all of them made use of performance assessments by external domain experts. Together, this set of studies allowed me to validate my hypotheses.

Table 3.1 provides an overview of the links between the themes, research questions, and component publications of the thesis.

The first research question (RQ1) is addressed in Chapter 4, which analyzes the potential of an existing adaptive interaction framework.

Publications I, VI, and VII fall under the first theme: conceptualizing and understanding information search. Publication I lays the groundwork for this research. It identifies several possible purposes of search and the state-of-the-art tools and methods applied for achieving these purposes. Publications VI and VII report on studies that compare the state of the art in exploratory search systems with a novel search system that provides an interactive visualization of the information space. The studies indicated that there is still room for improvement in user performance in exploration. These two publications answer the second research question (RQ2): does the information-seeker still need more support for exploration, even with existing tools and techniques?

Publication II and III address the second theme: modeling and predicting information search behaviors. In Publication II, a model for predicting the subjective specificity of search results in exploratory search is built in an attempt to answer the third research question (RQ3), on how one can predict the dynamic changes in the subjective specificity of information needs during exploratory search.

Publication III explores user behavior in both exploratory and lookup tasks to the end of building a classifier to distinguish these two types of search tasks. It thereby addresses the fourth research question (RQ4): can we distinguish exploratory search from lookup in the course of searching?

Table 3.1: Overview of the research themes, the research questions falling under each of them, and the component publications most directly addressing each research question.

Themes	Research questions	Publications
Conceptualizing and understanding information search	RQ1	This thesis as a whole
	RQ2	I, VI, VII
Modeling and predicting information search behaviors	RQ3	II
	RQ4	III
Providing real-time adaptive support for exploratory and lookup search tasks	RQ5	IV, V

Publications IV and V focus on the third theme: providing real-time adaptive support for exploratory and lookup search tasks. These publications deal with building of an adaptive IR system that predicts the search goal from user interactions, then dynamically changes the parameters used by the underlying IR algorithm. The target with these two papers was to provide real-time support for both exploratory and lookup tasks without altering the list-based interface familiar to users. These publications jointly address the fifth research question (RQ5).

The component publications together form a path to building information search systems that provide better support for both exploratory and lookup tasks.

Chapter 4

Formulating Information Search as Adaptive Interaction

This chapter presents a theoretical basis for formulation of information search strategies. It begins with an explanation of how exploratory and lookup search behaviors emerge from strategies aimed at maximizing utility in a given ecology with information processing bounds. This explanation is based on the existing framework of interaction strategies that I refer to herein as the AIF [123].

After providing an introduction to rational analysis and how the AIF has been developed on the basis of rational analysis, I offer a theoretical explanation of the search strategies that emerge in exploratory and lookup search activities, with reference to that framework. The chapter concludes with validation of the first two claims made in the thesis: 1) search strategies emerge as an adaptation to ecology, mechanism, and utility, in line with the AIF, and 2) the AIF explains why exploratory search is challenging.

4.1 Rational Analysis and the Adaptive Interaction Framework

Rational analysis is a theoretical concept in cognitive science that was introduced by Anderson to explain the function and purpose of cognitive processes [5]. It can be defined as an empirical method of explaining how and why a human cognitive system adapts. Based on the logic that humans optimize their behaviors to maximize the gains, it has been commonly applied to reasoning surrounding human behavior [38, 119]. As a higher-level concept, it has motivated two classes of model in the realm of information search: information foraging theory and economic models of search [17,124].

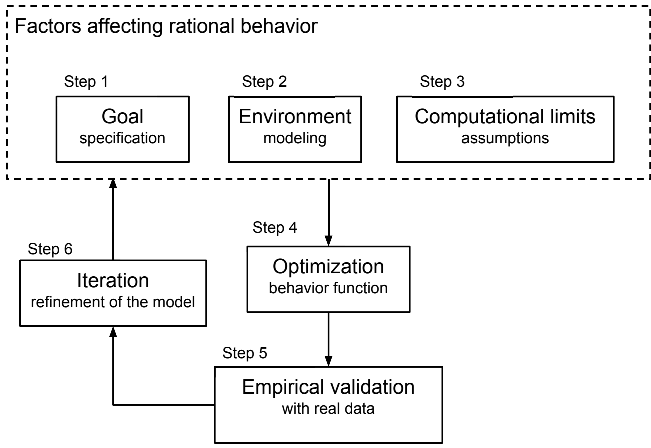

Figure 4.1: The main steps in the process of rational analysis. The first three involve specifying the three factors that affect rational behavior: goal, environment, and computational limits. In the fourth step, the optimal behavioral function optimization is derived. The fifth step is to validate the optimal function with real-world data, in what is referred to as empirical validation. Iteration through the process continues until the empirical validation confirms a good fit.

Rational analysis is performed as an iterative process [38]. Figure 4.1 shows the key steps in this process. It begins with the specification of three factors that affect the behavior: goal, environment, and computational limitations. In the first step in rational analysis, one specifies the goals that are being pursued with the cognitive system. Then, one builds a formal model of the environment wherein the system is operating. In the third step, assumptions are made about the computational limitations of the cognitive system. From these three factors one can derive the optimal behavioral function. Next, one empirically validates the optimal behavioral function with real data. To improve the accuracy of the behavioral function, the researcher needs to iterate through this process until the difference between the real-world data and the model's predictions is minimal.

Although in rational analysis the human is expected to select the optimal behavior, this is not what occurs in reality. Bounded rationality provides an explanation for this behavior: according to the principle of bounded rationality, humans make a compromise between the cognitive limitations and optimal behavior. Herbert Simon has explained the notion of bounded rationality thus: "rational behavior is shaped by scissors whose two blades are the structure of task environments and the computational capabilities of the actor" [138]. Human bounded rationality applies several mechanisms to deal with real-life complexity. Heuristic search is one scenario wherein there is a large space of possibilities to be explored [138]. If the task domain is poorly structured or of an unknown structure, we tend to settle for a solution that satisfies our expectations based on past experience. In this view, humans act as "satisficers," looking for a satisfactory solution rather than an optimal one [2].

The AIF extends bounded rationality analysis to the field of human–computer interaction [123]. It provides a logical explanation as to why users end up applying multiple strategies to interact with the same technology. One advantage of the AIF is that we do not need to make any assumptions about how users perform tasks. Rather, we can specify the interaction strategies as a machine-learning problem wherein the strategies emerge from the optimal policy [142].

The AIF possesses advantages over other quantitative models of information search behaviors. Information foraging theory [124] and economics-based models of search [17] are two other useful classes of model created in strivings to predict and explain the adaptive nature of human interaction with information. As already discussed in Chapter 2, both IFT and the economically based models mathematically quantify the actions available to the user in order to predict which actions the user would choose in order to maximize the rate at which the relevant information is acquired (see Section 2.4). However, both of these theories are concerned only about maximization of the rate of gain, whereas the user might find value in other factors. For example, there may be a user who has a daily routine of browsing the Web while commuting to work. Such a user is engaging in search for leisure rather than competing against time to maximize the information gain. The AIF also takes into account factors such as the cognitive and perceptual constraints and the user experience with the environment. Another benefit of the AIF is that it does not make any assumptions about the user. Therefore, I consider the AIF to be the most suitable framework for approaching the modeling of information search behaviors.

4.1.1 How the AIF Works

To understand how the AIF works, let us consider the example of searching for someone's contact number by using a cellular phone's contact list. A *strategy* consists of a sequence of interactions a user performs with the elements of the interface. For a given scenario, there could be many possible strategies in pursuit of the same goal. For instance, one way of approaching this task is to open the contact list and linearly scan it. Another strategy

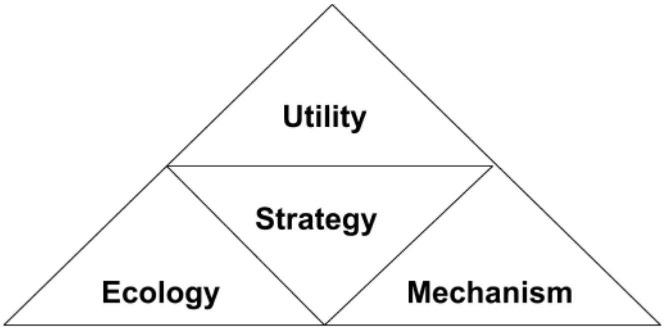

Figure 4.2: Elements of the adaptive interaction framework. The figure is taken from earlier work [123].

is typing the name of the person in a search of the contact list. In a third possible strategy, the user could open the phone's call-history log and scan through that. All of the possible strategies collectively form the *strategy space* of the user. The user's goal, knowledge of the environment, and computational limits determine what is the optimal strategy from within the strategy space. If the user is looking for a frequently contacted person, opening the phone's call log would be a good strategy. However, if the contact list is very short and organized alphabetically, scanning the list, including skipping to the relevant part of the alphabet, would be faster. These are some of the seemingly mundane strategies that a user could consider from the given strategy space. The rational user picks the most suitable strategy implicitly and quickly. The chosen strategy greatly affects the user's performance and satisfaction. With the AIF, it is possible to identify all the available strategies or the strategy space and also the optimal strategy that the user is likely to choose.

The main building blocks of the AIF are depicted in Figure 4.2. It shows that the interaction strategies positioned in the middle of the triangle are constrained by three factors, indicated in the three corners of the triangle: utility, ecology, and mechanism. These three factors lead us to quantifiable definitions of goal, environment, and computational limits in rational analysis.

Utility is an extension of the goal element in rational analysis and specifically refers to what the user finds value in. It is related to the utility maximization theory in microeconomics, wherein the goal is held to be maximization of the gain for the given budget (amount of money). In the example of searching the phone's contact list, the goal is to find the contact number of a specific person. If the user values time, the utility value lies in completing this task as rapidly as possible, or minimizing the time. If the time cost is too high, the user might consider the action to be of negative utility and hence try to avoid it.

Ecology pertains to the statistical structure of the environment as experienced by an individual. It involves both the immediate local task environment and the environment that the user has become familiar with over the course of a lifetime. For example, if the user knows from experience that the contact list is alphabetically ordered by last name and the person being looked for has a last name that starts with "a," the top of the contact list is the natural place to start the search. It is important to note that ecology is considered in relation to user experience or knowledge of the environment, rather than the true structure of the environment. Although the actual structure of the environment remains the same for all users, independently of the task (unless the system is adaptive), the ecology depends on the user, not just the task.

Mechanism here refers to human capabilities. It includes cognitive and perceptual limits in the human information processing system, such as the capacity and duration of human working memory and the latencies in motor movements. In our example case, parafoveal acuity and the time that it takes to fixate on an item, move the eyes, and read the name of a contact are some of the constraints. According to the AIF, the *strategy space* can be determined from the three components. The optimal strategy is the behavioral policy that yields the best gain in light of these: the ecology, utility, and mechanism. With the aid of the AIF and the principle of rationality, we can predict the optimal strategy or the policy that a rational user would select, which usually determines the sequence of actions performed when the user interacts with a technology.

In summary, the AIF can be used to model how the user interacts with a system by specifying the utility, ecology, and mechanism factors. One can determine the strategy space and the optimal strategy by solving a machine-learning problem. In Chapter 5, I will discuss how I trained an ML model to predict the optimal strategy in an exploratory search scenario using the AIF. One of the important contributions of the AIF is that it explains why users follow different strategies to interact with the same technology.

4.2 Lookup Search as Adaptive Interaction

The adaptive nature of humans in an information search process can be explained with the AIF. In this section, I explain the emergence of common lookup search strategies as adaptive interaction by considering the seemingly mundane lookup search task of answering fact-related queries such as "what is the predicted dollar-to-euro exchange rate?" Figure 4.3 illustrates this scenario as adaptive interaction.

In the latter scenario, utility is subjective and depends on what the information-seeker finds value in. It might be finding the most accurate answer or completing the task as soon as possible. If we consider a user who is involved in trading currency, accuracy of the answer might be the most important objective, but if the user is only interested in finding an approximate answer, then reasonably accurate information findable within the shortest time might be what is desired. With lookup tasks, in general, success in the task is measured in terms of task completion time [14]. Accordingly, Figure 4.3 represents utility as finding a reasonable answer as quickly as possible: minimizing the task completion time.

In this scenario, the ecology element is the user's familiarity with the statistical distribution of relevant answers on the search engine's results page. On account of his or her experience, the user may well expect the results to be ranked in descending order of relevance to the query issued. In Figure 4.3, I illustrate the ecology distribution as a graph of user-perceived relevance of the documents plotted against document rank on the SERP.

The mechanism could be influenced by the time it takes the user to read and comprehend each item in the set of search results. The structure of the SERP too could affect the mechanism. For example, if the user can see the answer on the very first SERP, it may take less time to process the result items. On the other hand, if the user has to scroll through the SERP to find the answers, the time cost is greater. A few examples of factors that contribute to the mechanism are given in the figure.

We could predict possible strategies on the basis of AIF. There are several actions involved in a strategy, such as formulating a search query, scanning a result snippet, following a link, reading a document, judging document relevance, and terminating the search session, as indicated in Figure 4.3. The user executes these actions with different probabilities; the likelihood of a transition between any two of the actions is referred to as the transition probability. The AIF aids in finding these probabilities by solving a reinforcement learning (RL) problem. In RL, a software agent is trained to perform actions in an environment so as to maximize the reward gained from each action (Subsection 5.3.1 provides a more detailed description of reinforcement learning) [142]. Thus we can predict the strategy space and provide a logical explanation addressing why a rational user would favor one particular strategy in preference to all other possible strategies.

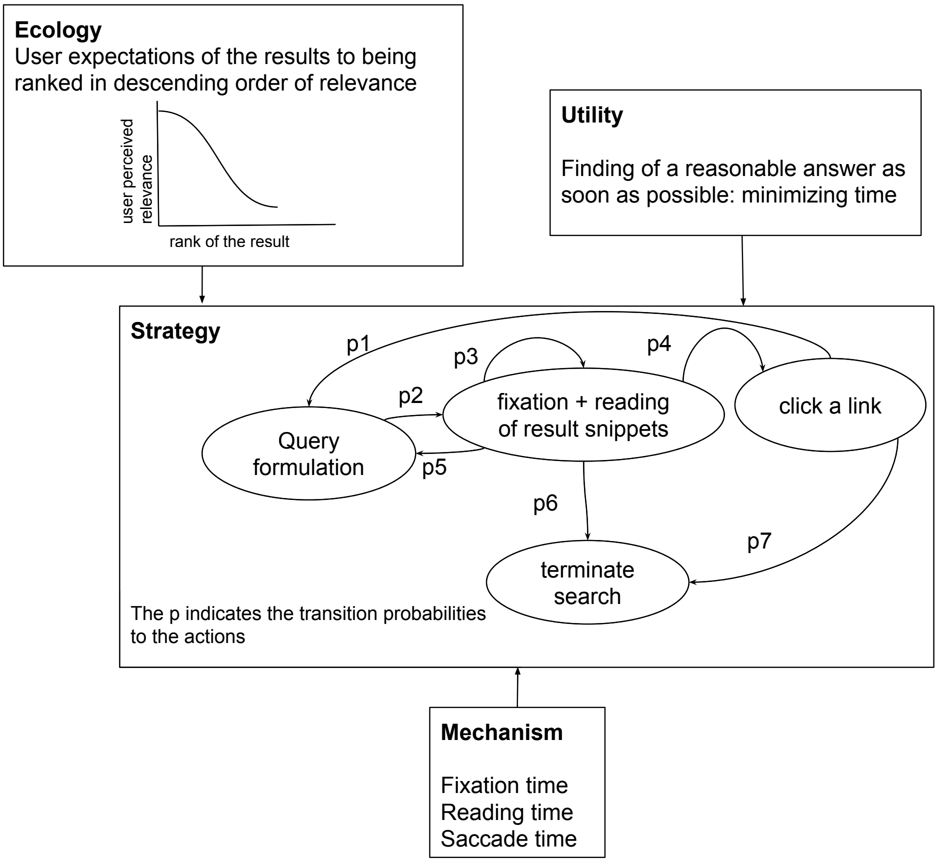

Figure 4.3: Lookup search as adaptive interaction. The ecology shows that, operating from past experience, the user expects the results to be perfectly ranked by relevance. A hypothetical ecology function is graphically represented in the figure. Utility is a function of time. Some cognitive and perceptual constraints included in the mechanism are represented in here. The strategy space is represented as a sequence of actions, where p denotes the transition probabilities for changing between actions. This representation of the strategy space is one possible approach to understanding strategy (adapted from the work of [19]).

4.3 Exploratory Search as Adaptive Interaction

There are various distinct characteristics of user interaction strategies in exploratory search, which were listed in Table 2.1. The AIF allows interpretation of these strategies as an adaptation to ecology, utility, and mechanism that is specific to exploration. Figure 4.4 depicts exploratory search as adaptation with respect to the AIF. For this analysis, let us consider a scenario wherein the user is trying to learn about a less familiar search topic.

In an exploratory search, the ecology can be expressed as the perceived relevance of the results. The user's familiarity with the information space in exploratory tasks is low [156], so the user has poor knowledge of the statistical distribution of relevant results. Other factors that contribute to the ecology are uncertainty surrounding the search queries made and diversity of search goals. If the search query does not articulate the actual information need, there is a mismatch between the IR-system-interpreted relevance and user expectations. When there are multiple search goals, the user is likely to find many results relevant. For these reasons, the statistical structure of the environment as experienced by the user would follow a less predictable pattern, with many peaks, as depicted in Figure 4.4. The figure presents the distribution of the user-perceived relevance of the search results by result rank. As the figure indicates, this distribution is less predictable.

Utility depends on the user's interpretation of value. As discussed in the background chapter, with exploratory tasks, the search goal is imprecise and open-ended. Also, the information needs are complex and involve answering "why," "what," and "how" questions [139] for the purpose of learning or investigation [107]. For such tasks, there is no single relevant document as in lookup tasks; rather, many documents are relevant, and these differ in their levels of relevance. To cope with such complex information needs in exploration, users seek documents that provide high information content [124]. In exploratory search, utility can be expressed in terms of the expected information gain.

The mechanism would involve document-reading and comprehension time—the time needed for understanding the content. In comparison to lookup tasks, reading and comprehension times are highly significant in exploratory tasks, because the user needs more time to understand and assess the relevance of a document when performing searches in an unfamiliar domain [157]. Existing knowledge of the search domain and experience with the IR system would be expected to influence query formulation skills. Factors related to the user's memory also have an impact on ability to recognize relevant documents, since exploratory searches are sometimes

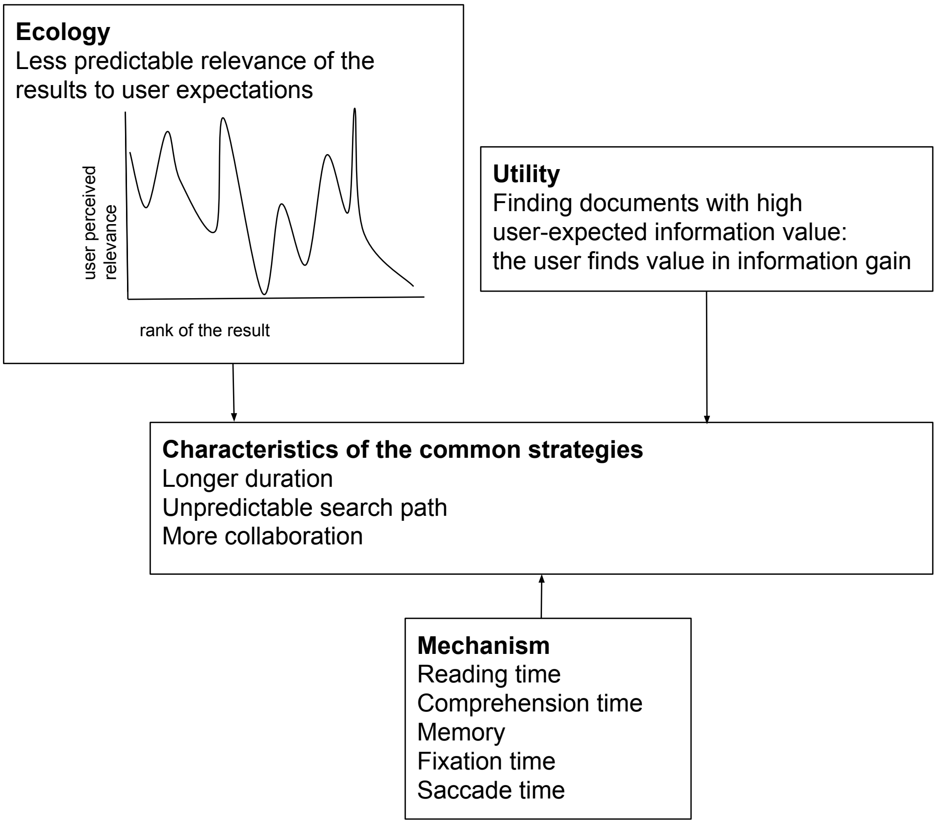

Figure 4.4: Exploratory search as adaptive interaction. The user-perceived relevance is less predictable—indicated by an ecology function with several peaks. Utility can be expressed as a function of information gain. In addition to reading, fixation, and saccade times in lookup tasks, the mechanism in exploratory tasks can be influenced by comprehension time and memory. They contribute to the three commonly observed characteristics of exploratory search strategies: greater duration, unpredictable search path, and more extensive collaboration.

conducted over a long span of time [157]. For example, a user with a good memory may remember relevant keywords discovered in previous sessions, while another user would need more time and assistance before recognizing a previously encountered keyword. All of these factors contribute to the information processing time in exploratory searches, as is indicated in Figure 4.4. Upon consideration of these factors, it is possible to derive the first claim.

4.3.1 Claim 1

Claim 1 states that exploratory search strategies emerging in adaptation to AIF. Ecology, utility, and mechanism together explain the three well-known characteristics of exploratory search strategies: longer duration, unpredictable search path, and more extensive collaboration, as indicated in Figure 4.4.

Duration: Exploratory tasks extend across multiple query iterations and sometime continue over numerous search sessions [157]. Lookup searches, which follow a "query and response" retrieval paradigm, take much less time to complete than exploratory searches do [139,163]. This phenomenon could be explained in terms of the differences in utility, ecology, and mechanism between exploratory and lookup tasks.

As figures 4.3 and 4.4 indicate, the ecology is less predictable in cases of exploratory tasks. This implies that the user has to devote more time to understanding the information space. The utility in exploration lies in finding documents that entail high information gain [125], rather than in completing the search task as soon as possible, as in lookup searches. With this utility framing, there is less time pressure, and those who embark on exploration generally have a larger time budget allocated for it [157]. In addition to the reading, fixation, and saccade times involved in the mechanism, which are common to all searches, exploratory searches feature a significant impact from additional factors, such as comprehension time and memory. These factors result in more time being used for processing of search results. If the user is a novice to the search domain, comprehending, reading, and truly understanding the documents requires more time [29]. Once users identify new keywords in the domain, typically they reformulate the queries with the newly recognized terminology [157]. In this connection, the user's memory has a considerable influence—if the user remembers newly encountered keywords, there is a higher probability of formulation a more precise query.

Search path: As was discussed in Chapter 2, exploratory searches display no clearly identifiable search path that leads to the desired results.

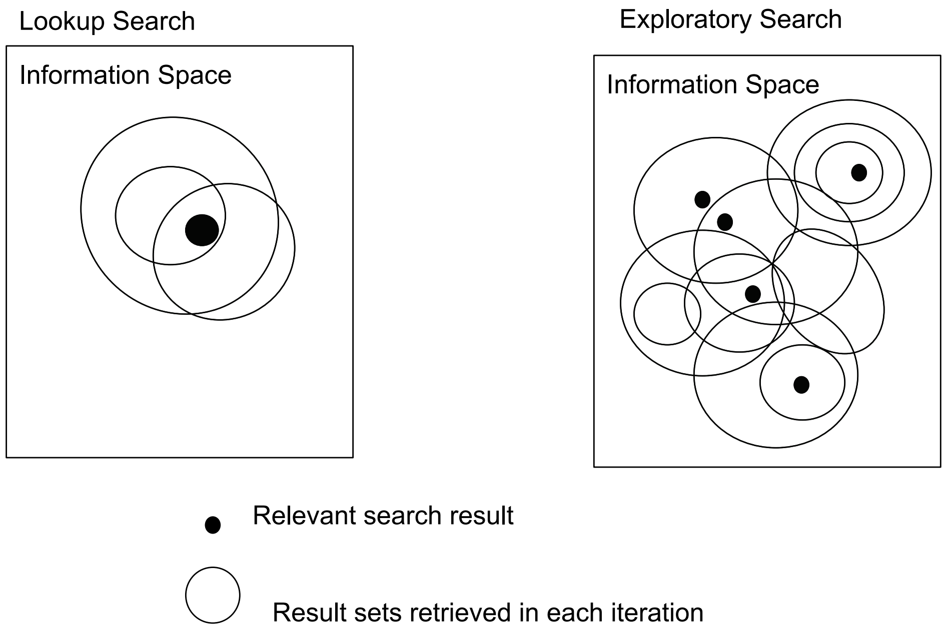

Figure 4.5: Search path and how users travel in the information space in lookup versus exploratory search tasks. The figure is adopted from [157]. The figure is related to the peaks in the ecology function indicated in figures 4.4 and 4.3. Many peaks in the ecology of exploratory searches correspond to the user having many patches of relevant information, indicated in the figure by many result sets with many relevant items. In lookup tasks, there will be only one or two relevant items.

Some literature refers to this phenomenon as high objective complexity [32]. Users traverse a much larger area of the information space and discover many relevant results (as indicated by the numerous peaks in the ecology in Figure 4.4). With lookup tasks, the search process is more straightforward and involves only a few iterations, which cover only the relevant area of the information space, in what is referred to as an iterative query-refinement strategy [22]. In typical lookup searches, there is a single highly relevant document with which the user terminates the search session. Figure 4.5 illustrates the search paths in the two main categories of tasks.

One key contributor to these differences in search paths between the two categories of tasks is the lack of user familiarity with the statistical structure of the information space, ecology, in exploratory tasks. Many documents exist, with varying degrees of relevance, so the user is uncertain about the way to approach accessing them. This leads to formulation of vague search queries that cover a broad swath of the information space and iterative reformulation of queries for examination of other parts of the space. The process is also referred to as sensemaking [40]. Another factor that motivates this behavior is the less constrained time budget for exploratory tasks—the view of utility. Unlike in lookup tasks, wherein the utility lies in finding the most relevant answer as swiftly as possible or minimizing the task completion time [14], in exploration the user attempts to get an understanding of the search domain as a first step, so as to be able to judge the relevance of the results accurately. Exploration is a highly dynamic process wherein uncertainty decreases over time and the scope of the information space is narrowed to more specific areas. In the end, the ecology distribution or user understanding of the distribution of the search space starts to follow a more predictable pattern, as shown in Figure 4.3, rather than a random distribution as shown in Figure 4.4. We could expect the search path in exploratory tasks to take on a pattern similar to that in lookup searches when this occurs [122].

Collaboration: Exploratory searches feature a high probability of collaboration with other people [157]. Some of those people may be involved in setting the goals that prompt the exploration. For example, in the academic context, a professor might advise a student to conduct a literature review on a given topic. In this scenario, the professor is interested in the outcome of the search task. One example outside the academic context would be a group of friends who are planning a holiday together. These people are involved in pursuit of a common goal. Another reason to involve multiple people in exploration is to obtain help when the task becomes challenging. The first two reasons are inherent characteristics of exploratory search, but the latter can be explained with the AIF.

If the user is uncertain about the distribution of relevant results and less familiar with the search topic—and hence less able to formulate queries iteratively for investigation of various areas of the information space—then he or she might well end up seeking help from domain experts. The time it takes to come to the decision about consulting a domain expert or collaboration depends on the utility and may equate to when the expected information gain is not reached within the allocated time. The mechanism too could contribute to this phenomenon. For example, if there is a user who needs more time to comprehend the information, consulting a domain expert might enable that user to perform more swiftly. If we consider the patch model in rational analysis, an information-seeker would decide to quit using a search engine and move on to a different information patch when the rate of information gain becomes less than the average information gain [125]. In exploratory information search, this new patch could be another person.

In conclusion, the main characteristics of the search process presented in Table 2.1 that distinguish exploratory search from lookup (duration, search path, and collaboration) can be logically explained via the AIF, as strategies that emerge as an adaptation to ecology, utility, and mechanism.

4.3.2 Claim 2

Claim 2 states that the AIF explains why exploratory search is challenging. Exploratory tasks are usually considered to be open-ended, abstract, and poorly defined [107, 157]. As is indicated in Figure 4.4, these factors contribute to unpredictable ecology, utility oriented toward maximizing information gain in a less constrained time, and a mechanism entailing issues of comprehension speed in addition to other human information processing capacities. Consideration of these factors led to RQ2: *do information-seekers still need more support for exploration, even with existing tools and techniques?* An empirical study to investigate how users perform search and to identify whether exploratory search is indeed challenging (as the AIF indicates) was clearly warranted. At the same time, empirical investigation was necessary for answering RQ2. I conducted the study reported upon in Publication I to investigate these aspects [7]. In addition to that study, two further studies led to similar findings and the same conclusion. Here, I focus primarily on the study for Publication I because its main purpose was to answer RQ2. I briefly explain the findings from the other two studies to provide further evidence, which has been reported upon in Publication VI [133] and Publication VII [64].

Study for Publication I: An investigation of information search behavior

The study reported on in Publication I [7] investigated the information search behaviors linked to various search goals to verify that exploratory search is the most challenging search task, a conclusion pointed to by the exploratory search behavior analysis performed with the AIF. This was a two-phase study with the first phase involving in-depth qualitative case-study work with mixed data collection methods: interviews, user observation, and diary-type logs. A mixed-methods study was necessary for obtaining thorough understanding of current search practices. The second phase involved a survey that corroborated the findings from the first phase.

Table 4.1: Literature search purposes (extracted from the case studies) and their frequency of occurrence (mean (m) and standard deviation (SD)) and subjective difficulty as rated by survey respondents (on a seven-point Likert scale). The statistical significance of frequency and difficulty ratings in comparison between the search purposes were tested via the Friedman test and follow-up pairwise comparisons with the Wilcoxon test. Those search purposes that occur most frequently are marked with three stars ($***$), purposes in the second most frequent band are denoted by two stars ($**$), and the least frequent purposes are marked with one star($*$). The search purpose associated with the greatest difficulty is shown in boldface.

Search purpose	Frequency	Difficulty
***Staying up to date with research	$m = 5.86$, $SD = 1.09$	$m = 3.05$, $SD = 1.51$
** **Exploring unfamiliar topics**	$m = 4.82$, $SD = 1.49$	$m = 4.04$, $SD = 1.72$
** Collaborating	$m = 5.07$, $SD = 1.60$	$m = 2.26$, $SD = 1.27$
** Reviewing literature	$m = 4.37$, $SD = 2.00$	$m = 3.17$, $SD = 1.31$
* Preparing lectures	$m = 3.65$, $SD = 1.92$	$m = 3.20$, $SD = 1.07$
* Recommending material	$m = 3.56$, $SD = 1.94$	$m = 3.20$, $SD = 1.32$

Academics active in the computer science domain were selected for this investigation, because exploratory search is most common among academics and because scientists are the heaviest users of electronic literature [75], with computer scientists being known to be early adopters of the latest search technologies [143]. Since one's academic experience has a profound influence on information-seeking behavior [117], individuals with three distinct levels of experience were selected to participate: Ph.D. (with at least one year of research experience and with one or more publications), postdoctoral (with at least five years of research experience and with one or more publications, as the lead author), and senior researcher (with at least seven years of research experience and currently leading a group or supervising more than one student). Two representative researchers at each academic level were recruited for study, for six participants in all.

The first phase of the case studies involved interviews that were aimed at finding answers to three key questions: 1) what are the main purposes of scientific information-seeking, 2) what search methods and tools do information-seekers use, and 3) what factors influence the search strategies? Participants first recalled all their reasons for searching for scientific information. Then, they positioned themselves with respect to their most recent search activity and walked through all the steps they followed, including the purpose that motivated the search activity, the entry point of the search, the tools used, how they navigated through the results, and factors that influenced the search process. This was a semi-structured interview. Participants also informed us about other methods that they follow to find information and how various individual purposes affect their search strategy, the search tools and methods used (in the past and at the time of the study), and how their search practices evolved over time. All the interviews were voice-recorded and transcribed before analysis.

The second phase of the case studies involved user observation. After the interviews, the participants were instructed to inform us when they were searching for scientific information for a real purpose. We then visited their workstations and unobtrusively video-recorded the search session. The participants thought aloud while performing each step. This allowed us to understand the reasons behind the steps they took.

The third phase involved longitudinal diary-based studies with the same participants. The main content elicited with the diary logs was 1) the purpose of the search, 2) the steps followed and tools used in the search process, and 3) user satisfaction with the findings. Participants were instructed to make entries at the end of every scientific search activity. Diary entries were collected for three weeks from the date of the interview.

A survey-style questionnaire was constructed after analysis of the case-study data. The case studies indicated six purposes behind academic searches. The questionnaire comprised one section per search purpose, and each section presented questions about the frequency and difficulty of searching for information for that purpose and on the importance of various navigation and sorting methods and tools commonly applied for that purpose. There were also sections about collaboration and collecting background information. For all the questions that involved ratings, we provided a seven-point Likert scale (with 1 being the lowest rating and 7 the highest). The survey respondents were those in the computer science discipline who were writing their master's thesis, conducting Ph.D. research, involved in post-doctoral research, or working as senior researchers. In total, 76 survey responses were received. The breakdown is as follows: 10% (8) from master's students, 50% (38) from Ph.D researchers, 24% (18) from post-doctoral researchers, and 16% (12) from senior researchers. Respondents came from 11 individual countries. and 42% of them were female.

Findings from the investigation of information search behavior

As mentioned above, the case studies identified six common purposes for initiating literature search. Table 4.1 provides a list of these purposes, along with their frequency of occurrence and their difficulty according to the survey responses. Exploring unfamiliar topics was reported to be the most difficult and the second most frequently occurring search purpose. According to the correlation analysis, there is a significant positive correlation between the frequency of exploring unfamiliar topics and its difficulty. This indicates that even with more practice, researchers still find exploratory search to be challenging. Also, there is a significant negative correlation between how well-established the literature is and difficulty in exploring unfamiliar topics. This indicates that researchers dealing with less well-established research areas find it even more difficult to explore unfamiliar topics. From these results we can conclude that exploring unfamiliar research areas is the search purpose that involves the most difficulty. Even with the invention of various tools to support exploration, users still find it to be a challenging exercise. Finally, this study empirically validates the claim I made on the basis of the AIF analysis with regard to the complex nature of exploratory search strategies.

Additional studies

Publication VI [133] and Publication VII [64] report on two additional studies, wherein users' performance of exploratory tasks with one of the most popular search engines for scientific exploration (Google Scholar) was compared with a system specially designed to support exploration (SciNet).

SciNet builds a model of the keywords extracted from documents to represent the search intentions of the user, then presents this model to the user. The visualization is interactive, meaning that the user can provide feedback and correct the system-predicted model. SciNet is designed to retrieve documents that are more relevant for exploratory search needs.

The primary objective for the studies was to investigate whether user performance in exploration can be improved via provision of special visualization support. These studies shed light on the issues raised in RQ2 by indicating whether there is still room to improve user performance in state-of-the-art search systems.

Two user studies, conducted to compare SciNet with two baselines, examined the retrieval performance (i.e., the quality of the results returned by the system in response to user interactions). The baseline system used in the first study was a within-system baseline setting in which users were only able to enter queries for SciNet, without benefiting from user modeling or interactive visualization—we refer to this condition as "simple SciNet." The second user study compared the quality of the user-retrieved information from SciNet with Google Scholar (the baseline system). In both studies, the users were placed in an exploratory search scenario with a task-based setting [78]. They were provided with a scenario describing information needs and asked to use the systems to obtain relevant information. Both studies were conducted in the laboratory in controlled settings, to restrict confounding factors such as social context and infrastructure (monitor size and computer speed) from interfering with the search session. User performance in the search sessions was judged via blinded assessment (i.e., the evaluators had no knowledge about the system used or the user) by expert researchers in the domain of the search topics. All documents that the participants retrieved during the studies were pooled, and the experts assessed them with respect to three properties: 1) relevance, whether the document is relevant or irrelevant to the search topic; 2) novelty, whether the document is related to a specific aspect of the topic rather than very well-known aspects; and 3) obviousness, whether the document is a very well-known article in the domain. These measurements were chosen because they have been concluded to be useful for evaluation of user performance in exploratory searches [43].

Findings from the additional studies

Considering the results of the first study, in which SciNet was compared to simple SciNet, we can suggest that the interactive visualization support in SciNet has a positive influence on the relevance and novelty of the retrieved documents. The experts' ratings indicate a statistically significant difference for precision, recall, and F-values between SciNet and simple SciNet with respect to relevance and novelty of the documents retrieved. No significant difference in obviousness ratings was found between documents retrieved from SciNet and from simple SciNet. These results provide evidence that search systems without added support for exploratory search are no longer sufficient to improve user performance of tasks of the type considered in the study.

In the second study, wherein SciNet was compared to Google Scholar, users were found to have retrieved significantly more relevant and novel documents with SciNet. This suggests that additional support could lead to improvements in users' performance of exploratory search tasks and that the popular existing tools need to be improved. There was no significant difference between the two systems with respect to obviousness ratings.

In conclusion, both studies shed light on the importance of providing additional support for exploratory search tasks. It was shown to be important to conduct further investigation of information searches, so as to identify techniques that can support exploratory and lookup searches equally well.

4.4 Discussion

In this chapter, I have introduced the AIF and explained the exploratory and lookup search strategies as an adaptation to ecology, utility, and mechanism in line with that framework, showing that the AIF provides a logical explanation for exploratory search behaviors and the complexity of exploratory search strategies. However, more effort is needed before we can understand how to differentiate between exploratory and lookup tasks, since the AIF does not point directly to an actual technique that enables the classification of tasks on the basis of search strategies. For the design of adaptive search systems, it is important to find features by means of which these two tasks can be distinguished. Another important element that demands further investigation is the dynamic nature of the exploratory search process. On account of the highly dynamic nature of exploratory tasks, utility and mechanism are subject to change. For example, at the outset in exploratory search, the user might be interested only in getting an overview of a topic, so the utility would lie in covering breadth for the topic rather than examining it in depth. Eventually the user might be interested in more in-depth information, however. In a similar vein, as the user gains more knowledge, reading speed and ability to recognize key terms should improve. This development affects the mechanism. The AIF needs to be populated with additional information on utility and mechanism adaptation functions before it can predict the evolving strategies in exploratory tasks. Furthermore, the framework as it stands does not explain how user satisfaction and overall performance of a task are affected by changes to ecology. For these reasons, the AIF on its own cannot be used to predict search strategies for all of the behavioral changes identified. Therefore, in addition to implementing the AIF, I have derived two other models to predict 1) how utility and mechanism evolve in exploratory search and 2) how one can separate between exploratory and lookup searches. Details of these are discussed in the next chapter. Also, I empirically verified that dynamic improvements made to the ecology enhance user performance and satisfaction, by developing an adaptive search system (Chapter 6).

Chapter 5

Modeling and Predicting Information Search

This chapter presents three attempts to build predictive models of information search behaviors. The first model was developed for predicting one of the dynamic parameters in exploration—subjective specificity of search results—from implicit user behaviors. The second model is a classifier that separates exploratory search behaviors from lookup behaviors while the user is still engaged in the search activity. Furthermore, I present the theoretical reasoning behind these two models in terms of the AIF. The third model involves computation of the optimal strategies in exploratory search (as discussed in Chapter 4). This is conceived of as a reinforcement learning problem in the AIF context and discussed accordingly.

5.1 Prediction of Dynamic Changes in Exploratory Search

We have discussed the dynamic nature of the exploratory type of search as one of the reasons for which it is considered difficult (see Section 2.2). We consider this problem more fully here. Since the user's knowledge, information needs, and search goals are subject to constant change throughout the search process, there are several layers to the issue of IR systems delivering results that satisfy the needs.

The objective for the second study was to build a formal model that predicts whether the search results are going to match the actual information need of the user from observable user interactions with SERPs. Moreover, it tackled the issue that whether the results for a query satisfy the user is highly subjective. This study answered the third research question (RQ3), which led towards my third claim: *the AIF can be used to model and explain how users adapt to search results that are either overly broad or excessively narrow in relation to their expectations.*

The process of exploratory search begins when a user has an interest in finding information on a topic of which he or she has little or no knowledge [157]. Typically, the user starts with vague queries and broad search terms. This allows obtaining cues about new keywords and repetitively reformulating queries with specific terms. Formulating good queries is challenging. When the user tries out various queries to get some idea about the topic, some queries might return documents that are overly specific, while other queries may lead to documents that are too broad to meet expectations. The objective with the study discussed here was to predict whether the search results would be too specific or too broad, relative to the actual information needs of the user.

In the study, the specificity of search results was modeled relative to the user's information need, which was referred to as subjective result specificity, or, for conciseness, "subjective specificity." The key objective was to envision IR systems that automatically detect the subjective specificity of a result so that they can effectively support the user's exploratory search. In particular, as the subjective specificity dictates, users will benefit from different types of support. For example, if the results are too broad, visualizations of the information space and guided tours should help the user to understand the new domain [64, 71]. If the results are instead too narrow, users might prefer introductory material explaining the new concepts, such as Wikipedia articles, or literature reviews [7]. Detection of subjective specificity should help IR systems support exploratory search through other techniques also, among them result clustering, keyword suggestion, and query expansion to determine whether the result set generated is too broad or narrow for the user.

I formalized a model that allows an IR system to infer subjective specificity from readily observable aspects of user behavior. That is, the model relies only on implicit click data and does not require any additional sensors, such as eye trackers [44]. Furthermore, the model is sensitive, in a predictable manner, to moderating factors such as prior search experience and in-session learning.

This section of the chapter provides an overview of the aforementioned model and the work carried out to evaluate that model. A rational explanation of why the model works, with the AIF as its underpinnings, is also provided. This model and the study are addressed in Publication II [11].

5.1.1 An Overview of the Model

The goal with the first model is to predict the effect of subjective specificity on exploratory information-seeking in which the user examines multiple search-engine result pages. The aim is to capture the iterative and evolving nature of search; that is, as the user explores a new domain, the search results become narrower and user knowledge expands.

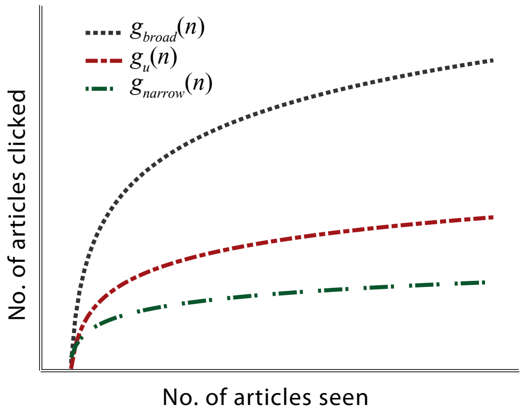

Figure 5.1: Hypothetical example of information gain as a function of the number of articles seen and that clicked on for exploration in more detail ("Seen–Clicked"). $g_u(n)$ is the user-specific effective information gain function. The model predicts that the gradient of this Seen–Clicked curve increases (e.g., $g_{broad}(n)$) when the results become more broad and that it decreases (e.g., $g_{narrow}(n)$) when the results become more specific than the current information need of the user would indicate. The figure is adapted from Publication II [11].

The model captures how *information gain* [124] in exploration behavior is affected by the subjective specificity. Here, information gain is defined as the number of search results that a user *clicks*, expressed as a function of the number of search results *seen* by the user. It is assumed that the information gain follows a natural logarithmic distribution. I adapt formulation work that has been carried out in information foraging theory [124] to predict how the slope of the information gain curve, the rate at which users click on results, changes when subjective specificity becomes low (broad) or high (narrow) with respect to the user-specific reference curve.

The key idea is that any given search result can differ greatly in the information content it holds for a certain user, depending on how well it matches that user's current information needs. Consider a user who is interested in learning about data-mining issuing a very broad query such as "data mining" at the outset. This query would yield broad results that include information about many areas, inviting the user to explore further. Consequently, the user would spend more time on every item [147], so we can expect a steeper slope for the information gain curve at this stage. Let us assume that in the process, this user encounters a new keyword, "subgroup discovery," and then formulates a very narrow query that uses this keyword. The resulting results retrieved would be too narrow and have very specific titles that make little sense to a novice, so the user probably would deem only a few of the items informative and worthy of further exploration. Here, the slope of the information gain curve is shallow.

The proposed model links two observable aspects of user exploratory behavior to what we call *effective information gain*. These are 1) the number of search results *seen* in a list and 2) the number of search results *clicked*. By using the word "clicked," I refer to the action of opening a link to a document in the search results for further investigation. I can now present the formal model devised for effective information gain (g) curve for a user (u) as a function of the number of result items seen, n, shown as $g_u(n)$ in Figure 5.1. This graph is referred to as the *Seen–Clicked* curve. Any gain function is affected by the objective relevance of the search results. In this case, when results are ranked by relevance, the function takes the shape of a diminishing returns curve as depicted in the figure.

The relationship between results seen and results clicked on can be described as a logarithmic regression model parameterized by λ and α:

$$g_u(n) = \lambda \ln(n) - \alpha \tag{1}$$

where n is the number of items seen so far in a result list (which is a positive integer the upper bound of which depends on the number of items on a SERP) and λ refers to the slope of this curve. α is a case-specific term that affects the maximum gain—it is determined by several factors, including subjective specificity and case-specific factors such as the search task and the maximum number of search results the user is expecting to receive. I make the assumption that when n is one item, α is -1 if the subjective specificity is broad and 0 otherwise. However, in reality n will be greater than 1. A logarithmic function is used to obtain the information gain, g, because this is the most natural foraging distribution [65] commonly used in models of human behavior [58]. With this model, we focus on the gradient of the gain function, λ, which depends on two parameters:

$$\lambda_u \ : \text{user-specific factor}$$

$$\lambda_r \ : \text{specificity factor for results}$$

The user-specific factor, λ_u, may depend on the user's experience with exploratory information-seeking or the search tool. For every user of a search tool, a distinct Seen–Clicked curve is defined by λ_u. The $g_u(n)$ curve plotted in Figure 5.1 shows an instance of a Seen–Clicked graph.

The results-specificity factor, λ_r, determines the effect of the subjective specificity on the gradient of the curve. For search results with high subjective specificity ("*narrow*"), the gradient of the curve reduces to a new effective gain function, as seen with the $g_{narrow}(n)$ curve in Figure 5.1. An instance of a Seen–Clicked curve for search results that have low subjective specificity, "*broad*," is indicated by $g_{broad}(n)$. While a single click carries little information about subjective specificity, the empirical data show that aggregate clicking behavior on a result page suffices for distinguishing among three levels (broad, intermediate, and narrow).

An IR system applying our model would monitor the clicking and viewing actions of a user in the course of a session. It would derive λ_u from the user's previous session, and throughout a given session it would derive λ_r from the user actions. In this connection, the gain function in Equation 1 can be expressed with a combination of λ_u and λ_r:

$$g_u(n) = \lambda_r \lambda_u \ln(n) - \alpha \tag{2}$$

A parameterized model predicts the subjective specificity of SERPs for the user and then compares the gradient of the Seen–Clicked graph based on the user's clicks on results of the current query with that of the user's baseline Seen–Clicked graph. Such a baseline graph can be constructed through measurement of the user's everyday interactions with the search tool. Then, if this user formulates a particular query to explore a research topic, the gradient of the new Seen–Clicked graph can be compared to the gradient of the user-specific baseline graph, so the system can predict whether prospective search results are going to be too narrow or too broad for the case-specific information need and adjust the system accordingly.

5.1.2 Estimation of Subjective Specificity in Exploration

I designed a study to validate this model. The work involved two sub-studies: During the first, I conducted a controlled laboratory study of exploratory search with a given set of search queries that had varying subjective specificity. In the second sub-study, I conducted free-form exploration wherein the participants explored topics of their choice. I will provide an overview of both sub-studies below. More details on them can be found in Publication II [11].

Sub-study I, with controlled queries

For the first sub-study, my subjects were 24 computer science students (master's- and doctoral-study-level) who were to search for scientific information on research topics that they were not very familiar with. The task for the participants was to collect scientific articles for the purpose of writing a scientific essay on a given topic. Six experts, in six distinct sub-fields of computer science, defined six unique tasks. The experts defined three search queries on each topic such that Google Scholar returned documents at a different level of specificity for each (one broad, one intermediate, and one narrow). Before the study, the participants were provided with a questionnaire to make sure that the subjective specificity of the queries corresponded with the participants' knowledge. An interface similar to Google Scholar was created to display the documents retrieved from Google via the three queries. Participants could scan and click on articles that they found useful for the given task. The tasks and the queries were randomized. By means of an eye tracker, the documents seen were logged. The documents clicked on were also logged. This enabled plotting the graph.

To confirm that, in accordance with our model, the gradients of the Seen–Clicked curves decrease as the subjective specificity of SERPs rises and that they follow a natural logarithmic distribution, we analyzed the overall distribution of the user information gain over information seen for the three types of SERP.

Sub-study II, with free exploration

In order to validate the model in a more natural setting, another study was conducted. It involved 10 computer science students exploring scientific articles in response to an authentic information need. Participants in this study were not involved in sub-study I. Four were M.Sc. students looking for scientific literature to inform their thesis projects. The other participants had just finished their M.Sc. studies and were exploring new research topics in preparation for making their Ph.D. research proposals. Google Scholar is the search tool they were all using in these tasks; therefore, an interface similar to that of Google Scholar was implemented to enable the participants to issue search queries and view results that were extracted from Google Scholar. A separate interface had to be created, because Google Scholar does not provide an application programming interface (API) to log user interactions. The arXiv database and its API were used to retrieve documents in response to the search queries. There were 40 documents displayed per result page, and these pages showed the same information as Google result snippets do. Every participant was allowed to conduct his or her natural exploration by means of our search interface for two hours. No restrictions were imposed on the search process, and the subjects could conduct searches in the same way as with Google Scholar: click on articles, read the articles opened, and take notes. The search queries issued, the results retrieved, and the articles clicked on were logged, with a timestamp for each. Experts in each search topic categorized the search results for every query as broad, intermediate, or narrow. The experts were either post-doctoral researchers or professors specializing in the search topic. Most of the experts (6/10) were supervisors of the participants so had an idea of the subjects' level of knowledge, which aided in prediction of the subjective specificity. For more control over the quality of the categorization, assessments were conducted by two experts in six out of the 10 cases. The Cohen's kappa test showed that there was substantial inter-annotator agreement (kappa coefficient = .67, p <.01).

5.1.3 Findings on Subjective Specificity

According to our model, the gradients of the curves should decrease with greater subjective specificity (or narrowness of the results), and they should follow a natural logarithmic distribution. To validate this hypothesis, we plotted curves for broad, intermediate, and narrow search results for each participant in sub-study I, with averaging over all the tasks they performed. By means of logarithmic regression, the model's parameters—gradient λ and case-specific term α—were computed for every participant. As expected, the logarithmic regression models for broad, intermediate, and narrow curves fit the data very well ($R^2 = 0.97$). Pairwise statistical analysis of the gradient of the gain curve between broad–intermediate, broad–narrow, and narrow–intermediate curves confirmed that when the subjective specificity increases, the gradient of the curve decreases. This suggests that the effective information gain declines with an increase in narrowness of the results. This is in line with the model.

For investigation of whether knowledge gain during searching, referred to as in-session learning, has an effect on the model, graphs were compared for the order in which the participants received the broad, intermediate, and narrow results. The results suggest that when the narrow results were considered after the broad ones, the gradients of the graphs for the narrow results were greater in comparison to when narrow results were considered before broad ones. This might be explicable in connection with in-session learning: when results gradually become narrower, the user is likely to make better use of the narrow results than when the change in focus level is in the opposite direction. On account of this behavior, when the narrow results were presented after the broad results, the number of results clicked by the user increased.

Also, the results suggest that our model is sensitive to the user's prior experience. When the user has more experience with exploratory search and with seeking scientific information, the gradient of the curve decreases, because there are experience-informed specific criteria as to the type of information deemed necessary.

Furthermore, a classifier was trained on the data, to evaluate the practical applicability of the model for predicting subjective specificity while the user is scanning the result list. Though the set of training data was small, the classifier (built with C4.5 decision trees) predicted the subjective specificity with 72.1% accuracy (area under the curve, AUC = 0.687) in classification of broad vs. narrow results.

Sub-study II confirmed that this model can be used to predict subjective specificity in natural exploratory searches. Proceeding from this result, I can propose that the model could be used to predict when a user actually needs help with narrow results.

5.1.4 Claim 3

Claim 3 states that the AIF can model and explain how users adapt to search results that are too broad/narrow for their expectations. The study, covered in Publication II, confirmed that information-seekers adapt their interaction strategies to the subjective specificity. If the results are too broad, then the user tends to follow a strategy that involves clicking many result items. If, on the other hand, the results are excessively narrow, the user would be expected to click only a few of them. Furthermore, the study showed the effect of in-session learning and prior experience on this behavior. Here, I provide a rational explanation for this adaptive behavior by considering the AIF.

In the framework's terms, overly broad or narrow results influence the ecology and mechanism. When search results are broader than what the user actually wants, the expected ecology distribution of the results changes—there are clearly too many results that are relevant to the topic, covering many sub-areas. In contrast, when the results are overly narrow, then the SERP contains many unfamiliar terms and the user ends up perceiving most of the results to be either irrelevant or less relevant. This too results in a different ecology distribution. Figure 5.2 illustrates these expected changes in ecology for overly broad and overly narrow results.

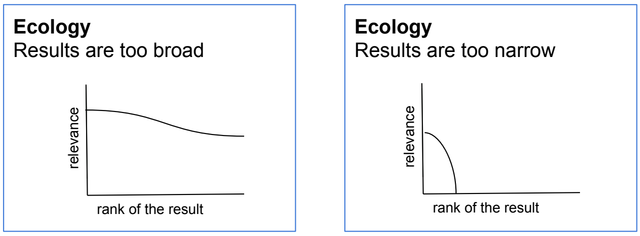

Figure 5.2: Hypothetical distribution of the results that the information-seeker perceives to be relevant—ecology—when the search query is overly broad (left) and overly narrow (right) in comparison to the actual information need.

We could expect the subjective specificity to affect the mechanism element of the AIF. If the user issues a query that produces overly broad and general results, many of them are likely to contain common and familiar keywords. For example, if a user who is exploring the topic of data-mining uses "data mining" as the query, most of the results are going to feature "data mining" in the title, as is shown by the screenshot of the top six results in the left pane of Figure 5.3. Less time is required for the user to read titles and comprehend the meaning when they contain familiar terms. If the user instead submits a query with a very specific keyword, it might return overly narrow results, containing many unfamiliar terms, as shown at the right in Figure 5.3. It takes more time for a user to read and comprehend the meaning of unfamiliar terms. Longer reading and comprehension times directly influence the mechanism. Accordingly, we would expect the mechanism to be affected by the subjective specificity of the search results.

Utility, however, is consistent, irrespective of the subjective specificity of the results, because the user's goal remains the same: to gain knowledge or explore an unfamiliar topic.

In a nutshell, when the results are too broad, the ecology distribution contains many results that appear relevant, and it should take less time to read and comprehend the results (mechanism). When the results are too narrow, however, only a few items are going to make sense to the user, and it will take a lot of time to read and comprehend the results. Given the same time budget and goal (utility), the information-seeker in the latter scenario (excessively broad subjective specificity) would follow a strategy that involves examining more items in the SERP than she examines when the subjective specificity is too narrow. This would lead to the information gain curve being steeper when the results are too broad than when they are too narrow. This explanation validates the third core claim.

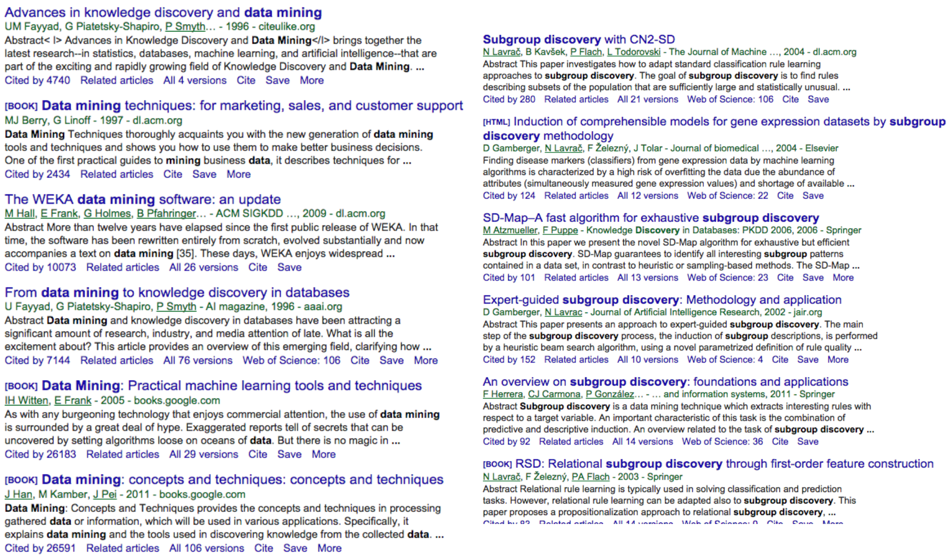

Figure 5.3: The top six results returned by Google Scholar for the search query "data mining" (left), which yields overly broad search results, and for the search query "subgroup discovery" (right), which produces overly narrow search results for an information-explorer who is unfamiliar with the domain.

5.2 The Model to Separate Exploratory Search from Lookup

One of the biggest challenges in designing search systems that support exploratory and lookup tasks equally well is that it is difficult to distinguish between these two main task classes early on in a search session. Therefore, a study was carried out with the objective of developing a model to distinguish exploratory from lookup tasks on the basis of implicit indicators of search behavior differences. That study answers my fourth research question (RQ4): *can we distinguish exploratory search from lookup in the course of searching?* This research question was investigated in Publication III [10].

It is difficult to tell exploratory and lookup search apart with IR systems. This is because a gap currently exists between our knowledge of exploratory search behaviors and the requirements imposed by IR system design. Thorough empirical analysis of exploratory and lookup activities within an IR environment is necessary before one can separate between the two categories. Moreover, to provide tailored and adaptive support, we should be able to predict the task type as early as possible. To this end, it is necessary to base the work on properties that can be measured from the first SERP onwards.

The primary objective with the study (which I denote as Study III) was to provide systemic and rigorous analysis of exploratory and lookup information search behaviors across several search activities. I approached this problem by initially operationalizing exploratory and lookup tasks, gleaning insights from an existing definition [100]. Then, I designed a user study to analyze information search behaviors, where information search behaviors are defined as the interactions between users and IR systems. The resulting study contributed systematic enumeration and quantification of behavioral variables that can be used to separate between the two classes of tasks. This section provides an overview of the operationalization, then briefly summarizes the study and the findings. Claim 4 follows from this work, and it is clear that the AIF explains the findings well.

5.2.1 Operationalization of Exploratory and Lookup Tasks

Discussed in Chapter 2, there are several attributes through which one can define exploratory tasks (see Subsection 2.1.1). To design tasks for research such as that in Study III, one has to consider the task description attributes: goal, search topic, and degree of uncertainty (see Table 2.1). I selected the goal and degree of uncertainty to operationalize exploratory and lookup tasks. One can measure the degree of uncertainty by considering the number of paths involved in the search process [32]; this is referred to as objective task complexity. Factors related to the search topic (the topic being less familiar and the search task being assigned or motivated by others) were considered in the recruitment for participation in the study.

Goal : In *exploratory* tasks, the search goal is imprecise and open-ended. That is, there is no *single* answer that fulfills the user's information needs and no clear criterion for when to end the search. Hence, the assessment of the relevance of results is not discrete. In *lookup* tasks, a precise search goal does exist. The search goal is reached by retrieving a finite set of relevant results, and the relevance of results can be assessed discretely.

Complexity : The objective complexity of a search task is commonly defined in terms of the number of paths involved in the search process [32]. This definition is intuitive and has been used in many studies [99, 104]. Clearly, for *exploratory* tasks, we cannot identify a single, direct path that leads to the desired results. Therefore, exploratory tasks show high complexity [157]. With *lookup* tasks, the search process is more straightforward and involves only a few steps—lookup tasks are typically of much lower complexity than exploratory search tasks.

The terms "core lookup" and "core exploratory" are used to refer to tasks that clearly possess the aforementioned characteristics. Core exploratory tasks display both high complexity of the search process and imprecise search goals, while core lookup tasks have low complexity and precise search goals. However, there are tasks with mixed characteristics; for instance, some lookup tasks have precise search goals but do not entail a straightforward search process. These are referred to as "borderline lookup" tasks [14, 104]. The other category of tasks, which have open-ended search goals but low complexity, is called "borderline exploratory" [116]. Study III explores how well IR systems can distinguish both core and borderline search tasks.

5.2.2 Study III—Prediction of Task Type from Information Search Behaviors

Study III involved designing exploratory and lookup search tasks with both core and borderline characteristics. Taking inspiration from Marchionini's task categorization (in Figure 2.1), I selected three exploratory tasks (knowledge acquisition, planning, and comparison) and three lookup tasks (navigational search, fact-finding, and question-answering). According to prior studies, of these three lookup task types, both fact-finding and navigational tasks display the core lookup characteristics, whereas question-answering tasks are identified by borderline characteristics [14]. Of the three exploratory tasks, knowledge acquisition and planning exhibit core exploratory characteristics [29, 116]. Comparison tasks were defined in such a way that they have borderline characteristics. After a review of information search behaviors identified in prior work, eight indicators of information search behaviors that one could expect to be both informative and at least somewhat easy to measure by means of an IR system were selected: query length, query duration, maximum scroll depth, cumulative clicks, proportion of browsing, duration of dwelling, task completion time, and gaze distribution.

The subjects, 32 researchers, searched for scientific articles satisfying tasks of each type, by using an interface we provided. The interface is very similar to Google Scholar and uses the arXiv.org API for obtaining results. No restrictions were imposed on either the search process or user interactions, but each task was limited to a 15-minute maximum.

5.2.3 Findings on Task Type Prediction (from Study III)

Statistical analysis of the eight indicators for information search behaviors revealed their power for discrimination of each task. Table 5.1 provides a summary of the overall statistical analysis conducted. The results empirically validated that IR systems can identify exploratory tasks within the first query session on the basis of various information search behaviors, including the length of the first query, scroll depth, duration of the first query iteration, proportion of browsing, dwell time, and task completion time. It is noteworthy that no significant difference among the tasks was found with respect to the gaze distribution.

To validate the applicability of our findings for a real-world IR system, classification experiments were performed, using state-of-the-art machine-learning methods. Apart from gaze distribution and task completion time, all of the behavioral indicators measured in the study were considered in this exercise. Gaze distribution was excluded because of the lack of significant difference referred to above and because this metric is not commonly available in IR systems. Task completion time was excluded because, naturally, it cannot be known while the user is still searching. According to a random forest classifier, the core exploratory tasks can be separated from the core lookup tasks with 85% accuracy (AUC = 0.859), but when borderline tasks are included, classification grows more difficult, and the accuracy for the tasks falls to 60.3% (AUC = 0.658). Further analysis confirmed that the borderline tasks show mixed characteristics—for example, in the question-answering tasks, the user behavior was very similar to that in exploratory tasks.

The data lead to elaboration on the original conceptual classification proposed by Marchionini [107], by pointing to information search behaviors that could aid in detecting with fine granularity when an exploratory task might take on aspects of lookup-associated behavior and *vice versa*. Proceeding from the set of information search behaviors proposed for real-world use, we can develop guidelines on how to distinguish between types of search tasks while the user is still performing the search. This should allow a search system to predict the task type for succeeding queries early on and adapt its support accordingly [137]. On the basis of these findings, a set of implications for possibly useful actions by IR systems can be derived: 1) adjusting the number of result items shown per SERP, 2) adjusting the length of the result snippets to match the task type, 3) using implicit relevance feedback techniques as the task type dictates, and 4) adjusting the exploration rate in keeping with task type.

5.2.4 Claim 4

Claim 4 states that exploratory and lookup searches can be distinguished from user interaction data and the AIF explains how. Chapter 4 presented a theory-oriented explanation as to why exploratory and lookup search behaviors are different, in line with the AIF. The findings from Study III empirically validated that it is indeed possible to separate between the two types of search tasks in an IR system while the user is still engaged in searching. Though the search tasks were controlled and not motivated by the information-seeker, the task context was created as naturally as possible in the given setting. Furthermore, the classification exercise confirms that an IR system can actually provide adaptive support for exploratory and lookup tasks.

With this finding, we could envision an information search system that is able to be used for both exploratory and lookup search tasks, equally well. These investigations have together answered my fourth research question.

5.3 Model of Rational Exploratory Search

The AIF provides a rational explanation for the exploratory behavioral strategies. In this section, I formulate exploratory search as a reinforcement learning problem and identify emerging behavioral strategies by finding the

Table 5.1: Predictive power of each feature, by task combination. The table shows how significantly the data for seven features differ between every combination of exploratory (knowledge acquisition, planning, or comparison) and lookup (fact-finding ("Fact"), navigation ("Nav."), and question-answering ("Q–A")) tasks. I used a Wilcoxon signed-rank test. Entries denoted by star symbols are significant after Bonferroni correction, with * for $p<.05$, ** for $p<.01$, and *** for $p<.001$.

Exploratory:	Knowledge acq.			Planning			Comparison		
Lookup:	Fact	Nav.	Q–A	Fact	Nav.	Q–A	Fact	Nav.	Q–A
					Query length				
p	**	***	.22	*	***	.53	***	.10	***
Z	2.97	4.4	1.22	1.72	4.07	.61	3.23	1.63	3.58
					Maximum scroll depth				
p	*	*	.83	**	***	.14	***	***	*
Z	2.21	1.9	.21	2.38	3.06	1.46	3.56	3.37	1.99
					Query duration				
p	***	***	***	.37	.97	.76	.14	.77	.91
Z	4.45	3.58	3.36	.89	.03	.30	1.48	.29	.12
					Proportion of browsing				
p	*	*	*	.39	.26	.57	.71	.79	.91
Z	-1.71	-1.82	-2.19	.85	1.12	.57	.37	.26	.11
					Duration of dwelling				
p	***	***	*	.47	.54	.82	.19	.54	.84
Z	3.80	3.42	2.89	.71	.60	−.21	1.30	.61	−.2
					Task completion time				
p	***	***	.12	***	***	.29	***	***	***
Z	4.3	3.8	1.5	4.4	3.8	1.06	4.5	4.2	3.4
					Cumulative clicks				
p	*	*	*	.61	.56	.69	.22	.19	.59
Z	2.57	2.39	2.34	.52	.58	.40	1.23	1.30	.54

optimal policy. This modeling approach allows computational validation of the first key claim made in this thesis, that exploratory search strategies emerge in adaptation to ecology, utility, and mechanism in line with the AIF (see Section 4.3). The section concludes with validation of the model via empirical findings from Study III.

5.3.1 Reinforcement Learning

Reinforcement learning is a computational approach used to simulate the well-known human behavioral strategy of learning from how the environment responds to our interactions with it [142]. The modern discipline of reinforcement learning emerged in the late 1980s and soon took two main paths: one stream of research, dealing with learning by trial and error, has its origins in the psychology of animal learning, and the other thread has to do with solving the optimal control problem by using value functions and dynamic programming. Reinforcement learning can be applied to situations that involve decision-making, such as a chess master making a move or a floor-cleaning robot deciding on its next move.

The reinforcement learning approach is focused on goal-oriented learning from interactions and on learning what to do to reach a given goal in a given situation. To implement it, we need to map the situation to actions and express the goal as maximization of a numerical reward function. Unlike in many other types of machine-learning approaches, the learner is not told which actions to choose in this scenario; instead, the learner discovers the best actions—the actions that yield the most reward—by trying them. Any method that is suitable for solving a real-world problem that can be expressed as an agent learning from the environment in order to reach a goal can be considered a reinforcement learning method. The agent applies experience to improve in performance over time. To solve a problem by using the reinforcement learning approach, one has to identify certain sub-elements of the learning system: a policy, a reward function, a value function, and (optionally) a model of the environment. *Policy* is related to the mapping between the perceived states of the environment and the actions to be taken when one is in those states. The reinforcement learning agent learns this policy during training. It is the core that determines the behavior of the agent. Most of the time, the policies are stochastic. The *reward function* determines the goal of the agent. It specifies the numerical value for mapping the perceived state to actions (in state–action pairs). The sole objective of the reinforcement learning agent is to maximize the cumulative reward. The *value function* specifies the best action to be taken in the long run. Though it is similar to the reward function, the reward function pertains to only the immediate gain rather than the long-term desirability of states. The *model of the environment* is the nature of the context that the agent is interacting with. It represents the behavior of the environment. When a certain state–action pair is given as input, the model of the environment predicts the resultant next state and next reward.

Below, in Subsection 5.3.2, I explain how I model the exploratory search scenario as a reinforcement learning problem for purposes of identifying the optimal policy by defining the policy, reward function, value function, and modeling of the environment.

5.3.2 An Overview of the Reinforcement Learning Model

The objective with the model developed at this stage was to predict the user interaction strategies with regard to the search interface from the standpoint of an analysis of what it is rational for a user to do. The model assumes that the user interactions are highly adaptive to the three factors in the AIF: 1) utility, or the goal for the task; 2) ecology, or the distribution of relevant results and the ranking order of relevant results; and 3) mechanism, or the cognitive and perceptual limits of the user. It takes the costs and rewards of each action into account in order to select behavior sequences that maximize the gain.

As has been discussed in Section 4.3, in exploratory search the utility is in finding documents with high information value; ecology would involve a noisy distribution of results perceived as relevant to the user; and the mechanism would involve reading time, comprehension time, memory, and gaze movement time (see Figure 4.4). If one is to implement the exploratory strategy computationally as an adaptation to these factors, it is necessary to represent the factors formally. The following key assumptions have been made for the formal representation of ecology, utility, mechanism, and strategy:

Utility: Utility is described as finding documents that the user perceives as relevant for the topic within the allotted time. Every document in the result list has some degree of relevance to the topic. Relevance is expressed as a real number between 0 and 1. Utility is represented as the information gain from interacting with the document.

$$\text{utility} = \begin{cases} 10{,}000 \times \text{relevance}, & \text{if action} = \text{read} \\ 30{,}000 \times \text{relevance}, & \text{if action} = \text{open} \\ 0, & \text{if action} = \text{fixate} \end{cases}$$

where relevance, again, is a real number between 0 and 1 that indicates how relevant the document interacted with is to the search topic. The values 10,000 and 30,000 are the weights assigned to the information gain for the read and open actions, respectively. The weight for information gain is selected in light of a similar model developed for menu search [41], and the weight for the open action is set through a trial-and-error approach. It is assumed that the information-seeker gains more information by opening a document than by merely reading its title. This is the reason for assigning greater weight to the open action. Finally, there is an assumption of no information gain for the fixation action. The action space is discussed in more detail later.

Ecology: Ecology is related to the user-perceived distribution of the search results. Here, I consider three distributions: the distribution of the documents relevant to the search topic, the ranking order of the relevant documents, and distribution by title length. I also assume the SERPs to have a fixed number of documents per page—referred to as the length of the result list. Relevant documents on a SERP are assumed to follow a power-law distribution [145]. Unlike in lookup tasks, wherein the search query represents the information need very well, in exploratory tasks we cannot assume the search results to be ranked perfectly by relevance to what the user wants to learn. Therefore, the ranking of the results follows a noisy descending order of relevance. The length of the document title affects the time cost of reading the title, so the user-perceived distribution for title length is considered in the ecology. Titles are assumed to follow a uniform random distribution between a minimum length of five words and a maximum of 15 [70]. Using this ecology representation, I generated SERPs for training the reinforcement learning agent.

Mechanism: Users perceive the relevance of each document fixated upon to the search topic. Perceived relevance is correlated with actual relevance, and the correlation coefficient depends on the type of action (fixation, reading, or opening). If the action is fixation on a document, the user cannot accurately perceive the actual relevance of the document: the correlation between the perceived relevance and actual relevance is very low. If the action involves reading the title, then the user can perceive the relevance more accurately: correlation between perceived and actual relevance is high. When the user decides to open the document subject to fixation, the perceived relevance is the same as the actual relevance. However, one must also consider that there is a time cost associated with every action. Fixation takes less time, and the average fixation duration for reading is 200–250 milliseconds [130]. Reading the entire title of a document uses more time, with that amount calculated by multiplying the mean reading speed by the number of words in the title. Finally, opening a document demands the most time—there is loading time and then document skimming time. I assumed the document skimming to involve reading the abstract of the document. I computed the mean skimming time by multiplying the fixation duration for reading by the number of words in the abstract. Memory capacity determines whether the user remembers a document that has already been interacted with. It is assumed that the user remembers the most relevant and least relevant item that he or she has read or opened.

Strategy: A reasonable strategy (or policy) for exploratory search is to optimize the information gain (ecology) while minimizing the time cost (mechanism) for a given ecology. I assume that the state space consists of the fixation position, the perceived relevance, past memory of the most and least relevant document that the user has already opened (clicked on) or read, and all the documents that the user has opened/clicked so far. Items opened / clicked on are added to the state vector because browsers, in general, highlight the links for items that the user has recently clicked on.

For clarity of understanding, let us imagine that a user with experience of scientific information-seeking is interested in acquiring new knowledge of a less familiar topic. After issuing the first query, the user gets the first SERP, illustrated at the bottom right in Figure 5.4, above, and this SERP shows 20 documents. The goal is to find as much information as possible about the topic that would help the user to gain more knowledge.

The user might approach this task by first fixating on the top document the SERP shows. After fixating on it, the user would perceive a relevance level of the document with respect to the topic, but that perception of relevance might not be very accurate, since the user had only a very quick (200 ms) look at the title of the document. The state space is updated on the basis of this action.

States: For a SERP with n documents per page, the state space can be represented as a vector V with five elements. Vector V is composed of one element for the position fixated upon, which indicates the rank of the item subject to fixation; one element for the perceived relevance of the

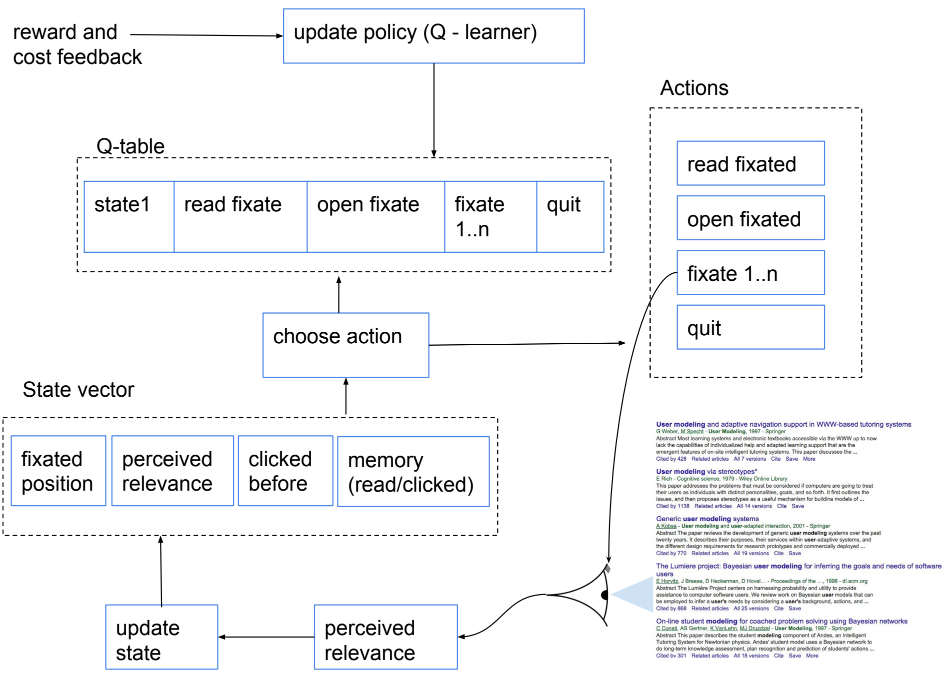

Figure 5.4: An overview of the adaptive exploratory search model adopted from [41].

item fixated upon; one element indicating whether the user has clicked on this item before or not; and two elements for the rank of the most relevant and least relevant item that the user has encountered so far. The fixated position is an integer representing one of the n rankings of items on the SERP [1...n]. Four levels are used for perceived relevance [Null, 0, 1, 2]. Clicked-before is a binary value [Null, 0, 1]. Memory has $2\times n+2$ possible values [Null, 1...n, Null, 1...n] (this holds the position of two items). All of the items in the state vector are initially null, before the user perform some action.

Actions: For each state, there are $n+3$ actions: n actions for fixating on a document n on the SERP, an action for reading the title of the item fixated upon, an action for opening the fixated-upon item by clicking the link, and an action to quit or terminate the search session.

After the first action, the user receives a reward for the action.

Reward: Reward is computed by combining the expected information gain formalized in utility terms and the costs related to the cognitive and perceptual constraints in the mechanism.

$$\text{reward} = \text{information gain - cost of action} \tag{5.1}$$

$$\text{information gain } = \text{utility} = \begin{cases} 10{,}000\times \text{ relevance}, & \text{if action} = \text{read} \\ 30{,}000\times \text{ relevance}, & \text{if action} = \text{open} \\ 0, & \text{if action} = \text{fixate} \end{cases}$$

$$\text{cost of action} = \begin{cases} \text{RS} \times \text{title\_length}, & \text{if action} = \text{read} \\ \text{BT} + \text{LT}, & \text{if action} = \text{open} \\ 200ms, & \text{if action} = \text{fixate} \end{cases}$$

where *RS* is the mean reading speed (in milliseconds per word), *title_length* is the number of words in the title, *BT* is the mean browsing time to explore an opened document, and *LT* is mean document loading time. Both *BT* and *LT* are given in milliseconds. The variable *relevance* is the actual relevance of the document interacted with for the search topic. Here the fixation duration is set to 200 ms and it is assumed that there is no information gain (information gain = 0) for fixating because the user cannot comprehend any information in so short a time [130].

Learning

This reinforcement learning problem was solved by means of a well-known algorithm for reinforcement learning, called Q-learning [142]. Q-learning obtained the value of each state–action combination with simulated SERPs. It maintains the learned state–action value, referred to as the q-value, in a table called the Q-table. The Q-table is a matrix in which every row corresponds to a unique state and the columns represents the actions. In the search scenario, there were $n + 3$ columns and the length of the Q-table (that is, the number of rows) grows in accordance with the number of unique states. The state–action values in the Q-table are updated incrementally (i.e., learned) as the reward is calculated, on the basis of the cost feedback and information gain from the chosen action, as expressed in the utility function. The next action is chosen on the basis of the available state–action values in the Q-table. Although I used Q-learning to solve this problem, any Markov decision process (MDP) solver that is guaranteed to converge to the optimal policy can be used to derive the rational adaptation [142].

The model was trained on 10 million trials, until there was no significant difference in the q-values in each of them. In each trial, the model was trained on SERPs constructed via random sampling from the ecological distribution of relevant documents and the title length distribution. Parameters required for calculation of the cost of actions, such as RS (reading speed), BT (browsing time), and LT (loading time) were computed from the data collected in Study III. Also, the distribution of the title lengths was constructed from the titles of the $6,000$ unique articles that the participants retrieved in that study. The model explored the action space by using greedy ϵ policy, meaning that the agent exploited the greedy/best action with a probability of 1 - ϵ and explored the actions randomly with probability ϵ. The optimal policy is the greedy policy given the q-values. The optimal policy was then applied in a further 10,000 trials of newly sampled SERPs for recording of the final strategies.

5.3.3 Computational Validation of Claim 1

Task completion time: All the results were obtained with application of the optimal strategy. Figure 5.5 indicates the task completion time for the reinforcement learning model and for our human users. As Section 4.3 has highlighted, information-seekers take longer for exploratory tasks. In Study III (which provided data for the validation), the participants were given a maximum duration of 900 seconds (15 minutes) for each task. The same time constraint was set for the model—in each trial, the estimated time cost for all chosen actions when added together gave us the duration calculated for the trial, and the trial was terminated if the duration rose to above 900 seconds. As Figure 5.5 indicates, both the model and the exploratory information-seeker required more time to complete the tasks; the lookup

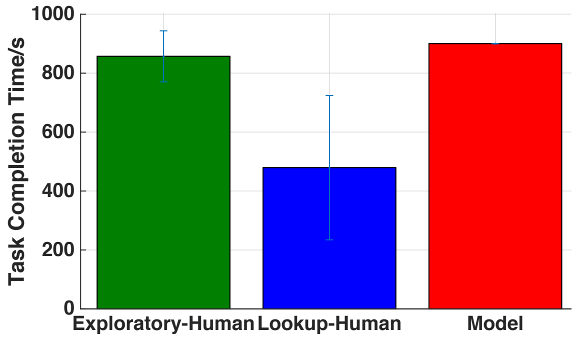

Figure 5.5: Exploratory task completion time (mean and standard deviation) for a human user and for the reinforcement learning model, compared with the lookup task completion time of human users. Note that both the model and human users take considerably longer to complete the exploratory task than the human user takes for a lookup task.

tasks were of considerably shorter duration. These data computationally validate one of the user strategies proposed in the first claim in light of the AIF: exploratory search tasks take longer than lookup search tasks.

Proportion of clicks: Figure 5.6 shows the mean percentage of documents opened for each rank of results on the SERP. When performing lookup tasks, users tend to open the document ranked at the top of the SERP the most. This strategy is expected, because, according to the ecology, the item at the top of the SERP is expected to be the most relevant document (see Figure 4.3). With exploratory tasks, however, both the model and the human user are less likely to click on or open documents.

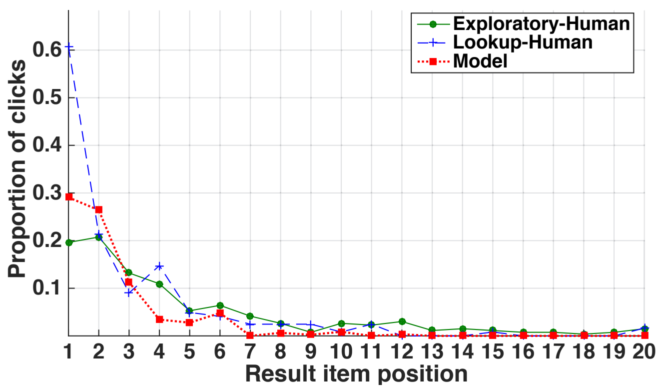

Figure 5.6: Proportion of clicks for each rank on the SERP.

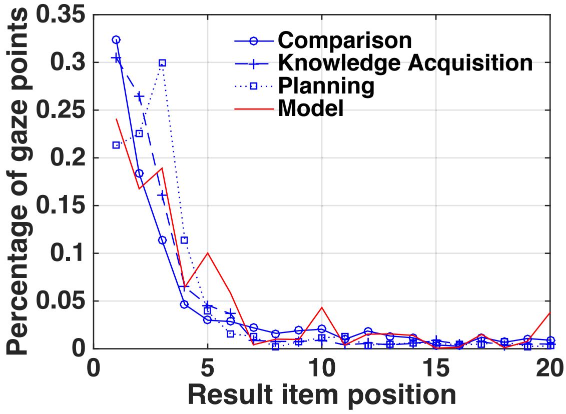

Figure 5.7: Proportion of gaze points for three distinct types of exploratory tasks (comparison, knowledge acquisition, and planning) and the reinforcement learning model.

This might be explained by the high cost of opening a less relevant document. Since users cannot accurately perceive the actual relevance of the documents by reading the titles, they more often refrain from clicking them. As Figure 5.6 shows, there is a clear match between the document clicking/opening strategies of humans and those of the model for exploratory search tasks.

Proportion of gaze points: Figure 5.7 indicates the proportion of gaze points or fixations for results with each ranking on the SERP for the reinforcement learning model and the equivalent for human users carrying out the three types of exploratory tasks examined in Study III—comparison, knowledge acquisition, and planning. This analysis also provides evidence that the model follows strategies comparable to those employed by actual human users performing exploratory tasks. However, for some rank positions the model does show slight deviation from the actual user behavior. This might be due to differences between actual users' memory and the assumed memory in the model. This element calls for further research for extension of the reinforcement learning model derived from the AIF to more accurately parameterize memory-related and other cognitive constraints affecting the user.

5.4 Discussion

This chapter has introduced three models that I developed to support information search. The first model is capable of predicting whether a SERP is either overly broad or overly narrow in comparison to what the user expects. With the second model, an approach is proposed for separation of exploration from lookup searches while the user is still engaged in the search session. The third model implements the AIF for exploratory search tasks by using reinforcement learning. The findings surrounding the models are exciting; we learned that it is possible to make educated guesses as to the search goal (exploratory or lookup), how well the results match the user needs (subjective specificity), and how the user would interact with the SERP when carrying out exploratory tasks.

These models have valuable implications for the design of IR systems. The models' detection of subjective specificity and discrimination between exploratory and lookup tasks can be exploited in the design of adaptive IR systems. The third model, which implements the AIF, is important for predicting how a user would interact with a given search interface for exploratory tasks. It could be used by search interface designers to evaluate various interface design options. For example, the model could predict how the gaze distribution, clicks on result items, and task completion time would change when the user interface includes other features. This research holds promise for fruitful developments in which an IR system could predict how new interface elements affect user performance without these actually having to be implemented.

Chapter 6

Real-Time Support for Exploratory Search

This chapter provides an overview of research conducted to develop a search system that provides real-time support for both exploratory and lookup tasks. I use the term "real-time support" to refer to information search systems that are able to predict the type of search task while the user is actively searching and use that prediction for dynamic adaptation of system features, to provide customized support for the task at hand. The work reported on here was motivated by the fifth research question (RQ5).

6.1 Parameterization of Exploration Rate

In recent years, reinforcement learning (RL) techniques have demonstrated great potential for supporting exploratory search in IR systems [129, 171]. The approach allows the system to trade off between exploitation (moving toward more specific topics) and exploration (presenting the user with alternative topics and thereby enabling the system to build the user model gradually without getting stuck in a local search space) [98]. Considering the benefits of RL techniques for both exploratory and lookup tasks, we selected an RL algorithm for the retrieval algorithm in the proposed system.

Parameterizing exploratory systems, however, is not trivial. Because of the inappropriately tuned exploration rate in IR systems that employ RL techniques, users often do not feel that the system is responsive to their information needs. Particularly when the search task is of the lookup type, users would usually prefer the IR system to exploit the search query issued and return the results best matching it. On the other hand, with exploratory tasks, in which even the users are uncertain about their queries, the search query should be taken with a grain of salt. Hence, the right balance between exploration and exploitation plays a crucial role in user performance of search tasks, of various types. Therefore, IR systems should trade off between exploration and exploitation on the basis of task type.

I conducted two studies to parameterize exploration rate (Study IV and Study V). Study IV represents an approach to finding the ideal balance between exploration and exploitation for exploratory search tasks. This approach was reported upon in Publication IV [8]. When this work was complete, an IR system was built to adjust the exploration rate dynamically in line with the task type. Study V empirically validated that real-time tuning of an exploration–exploitation parameter in an RL-based IR system improves users' performance of both exploratory and lookup tasks. That study has been reported upon in Publication V [9].

6.1.1 Study IV, on the Tradeoff between Exploration and Exploitation

An overview of the system

An IR system was designed with a simple list-based interface (see Figure 6.1) to retrieve 20 documents per query. With this system, the initial set of documents is ranked via the Okapi BM25 algorithm [141]. To see more documents, the user can indicate which document is of interest by clicking on the circle next to it, thereby assigning a relevance score of 1 to that document. After clicking on the Next button in the top right corner of the page, a new set of documents is displayed, with content based on the feedback provided by the user so far. Documents that do not receive explicit relevance feedback are assumed to merit a relevance score of 0. The document corpus indexed by the system consists of approximately one million documents obtained from the arXiv repository.

The RL algorithm LinRel [12] was used to allow the user to explore the document space. This algorithm begins with a matrix D in which each row d_i is a tf–idf (term frequency – inverse document frequency) feature vector representation of the initial set of documents presented. The user-provided relevance feedback on retrieved documents up to time t is represented as a columnar vector $r = (r_1, r_2 ... r_t)^\top$. The expected relevance r_i of a document d_i is expressed as $\mathbb{E}[r_i] = d_i \cdot w$, where vector w is estimated from user feedback. LinRel estimates $\hat{w}$ by solving $r = D \cdot w$ and produces an estimated relevance score for each d_i as $\hat{r_i} = d_i \cdot \hat{w}$.

In order to deal with the exploration–exploitation tradeoff, the algorithm does not automatically present those documents with the highest

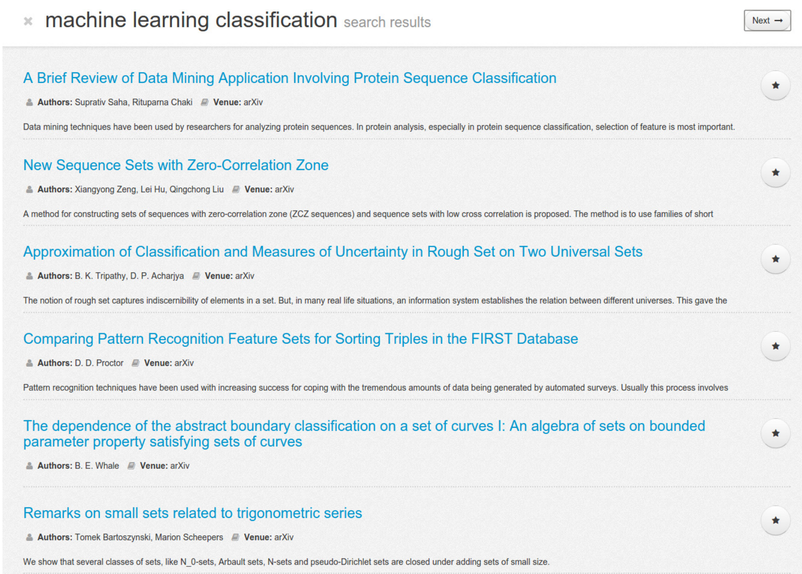

Figure 6.1: A screenshot of the interface developed. The search query is displayed at the top. A document that is of interest can be selected by clicking the icon next to it, and the user can proceed to the next iteration by clicking the Next button. Figure taken from earlier work [8].

relevance score $\hat{r}_i$. Instead, it retrieves the documents with the highest upper confidence bound for relevance. Therefore, if σ_i is the upper bound on the standard deviation for relevance estimate $\hat{r}_i$, the upper confidence bound for document d_i is calculated as $r_i + \gamma\sigma_i$, where $\gamma > 0$ is a constant used to adjust the confidence level for the upper confidence bound. In each iteration, LinRel calculates $s_i = d_i \cdot (D^\top \cdot D + \lambda I)^{-1} D^\top$, where λ is a regularization parameter that is set to 1 if each of the feature vectors sums to 1 (following previous work [12]) and the documents that maximize $s_i \cdot r + \frac{\gamma}{2}\|s_i\|$ are selected for presentation. The first term, $s_i \cdot r$, effectively ranks all the documents for their similarity to the documents the user has selected thus far; it thereby narrows the search space (for exploitation). The second term, $\frac{\gamma}{2}\|s_i\|$, ensures that the user is presented with a set of results that has greater variety. The parameter γ deals with the exploration–exploitation tradeoff. The higher the value of γ, the broader, or more tuned for exploration, the results are.

The user study

The goals for the user study were to investigate how differences in exploration rate affect 1) the number of documents selected, 2) subjective perception, and 3) overall performance. Ten academics each performed five exploratory search tasks, with five unfamiliar search topics, for the purpose of writing a scientific essay, using our system. Simulations were run to identify a possible set of exploration rates to test in the user study. Five distinct exploration rates were selected (γ) through the simulations: 0.0, 0.2, 0.5, 1.0, and 2.0. With these exploration rates, 0, 1, 3, 5, and 9 exploratory documents were retrieved, respectively. This means that when γ was set to 0.0, the SERP contained documents that had a perfect ranking with respect to the query and there were no exploratory documents. On the other hand, when γ was set to 2.0, nine of the 20 documents on the SERP were exploratory. More details about the simulations can be found in Publication IV [8]. Every participant performed one search task with each of the exploration rates.

6.1.2 Findings of the Exploration Rate Study

Multiple regression analysis was performed on the user study data (Study IV), to predict the number of relevant documents selected by the user for the various exploration rates. Results of this analysis showed that an exploration rate of 1, which roughly corresponds to a quarter of the documents presented being results of exploration, leads to the highest number of documents being perceived as relevant by the user. This finding was confirmed by qualitative analysis of user-perceived satisfaction with the search results and by the task performance results. The results together provide valuable insight into optimization of RL-based exploratory IR systems and have implications for design of self-adapting IR systems.

6.1.3 Study V and the Adaptive Information Retrieval System

The objective in Study V was empirical analysis of an adaptive search system that built on the output from studies III and IV: 1) Study III's classifier that recognizes the task type (lookup vs. exploratory) while a user is searching and 2) the reinforcement-learning-based search engine from Study IV, which adjusts the balance between exploration and exploitation accordingly in the ranking of documents. The adaptive search system developed in this study allows supporting both exploratory and lookup tasks surreptitiously without departing from the familiar list-based interface. The search results have more diversity when users are exploring, and more precise results are returned for lookup tasks. This study and the findings are discussed more fully in Publication V [9].

An overview of the system

Figure 6.2 provides an overview of the system. In essence, this system's features are very similar to the features of the system used in Study IV: after typing in the search query, the user is presented with 20 documents, and the interface is the same as that presented in Figure 6.1. However, there are two main differences—this system includes a classifier to predict the task type, and the retrieval algorithm utilizes an adaptive exploration rate rather than a static/fixed one. When wishing to explore more documents, the user can click the Next button in the upper right-hand corner of the page. Then, the classifier predicts the task type—exploratory or lookup—by considering the user's browsing behavior with the results in the first iteration. The classifier was built in keeping with the approach proposed in connection with Study III, and it was trained on the same data. If the classifier predicts the task to be exploratory, the exploration rate is set to 1 (in response to the findings from Study IV), and if the task is of the lookup type, the exploration rate is set to 0.

The user study

The purpose behind the user study was empirical validation of the hypothesis that an adaptive search system, dynamically adjusting the exploration rate to match the task type, improves user performance in comparison to traditional search systems, which are designed to support only lookup tasks. A system with the same interface (see Figure 6.2) and the same retrieval algorithm but with the exploration rate set to a static value of 0 was used as the baseline. This setting was selected for the baseline system because most existing search systems are similarly tuned to exploit the search query with no adaptation to task type [107].

The study involved 18 researchers performing two search tasks, one exploratory and one lookup, with each system. With this study, our goal was to create naturalistic search tasks. To this end, instead of imposing search topics for exploration, we asked the participants to come up with less familiar search topics that they were interested in learning about. The lookup tasks were designed in accordance with a typical known-item search scenario, in which participants were given target articles two hours prior to the study to skim through and then were asked to find these articles.

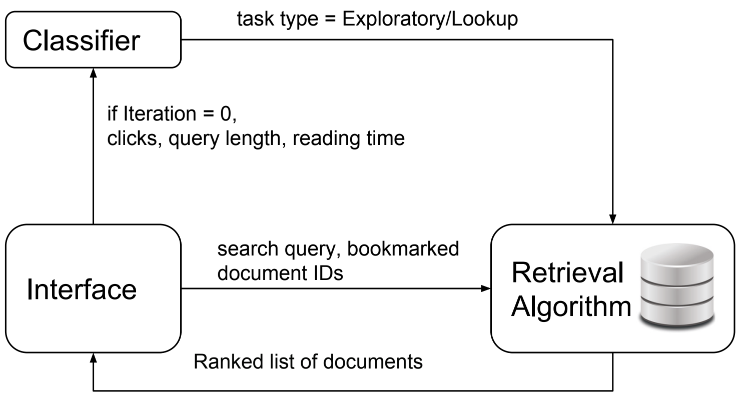

Figure 6.2: An overview of the system designed. The user types in a query and investigates the first page of search results. From the user interaction with the search results displayed on the first page, the search is classified as either lookup or exploratory. The result from the classifier is passed on to the search engine, where the exploration rate for the retrieval algorithm is set in accordance with the search task. Once set, the exploration rate is kept constant for the remainder of the search session. The figure is taken from previous work [9].

The system logged all the interactions initiated by the user and recorded the classifier-predicted and actual task type. We also collected qualitative feedback on the search system by means of interviews. Using a binary scale, an external expert reviewer with specialist knowledge of the search topics rated the relevance of all articles returned by both systems for the exploratory tasks.

6.1.4 Findings from the Adaptive IR System Study

The user-provided ratings (in Study V) for the usefulness of the selected articles suggest that for exploratory tasks our system with an adaptive exploration rate retrieved more useful documents than the baseline system did. The qualitative feedback provided during interviews further confirmed that users perceive the full system to be better for exploratory search tasks. This finding shows that users are sensitive to the changes made in the exploration parameter and a higher exploration rate is necessary for exploratory search tasks. The classifier classified 81% of the tasks accurately. This is an important finding, given that the exploratory search tasks in this study were truly motivated by the participants. In addition, the results reveal an interesting difference between the number of results the users bookmarked as relevant and how many the expert judged to be relevant. The expert rater found no difference between the search results returned by the two systems; however, users tended to bookmark more documents as relevant in initial iterations when using the adaptive system than in the baseline condition. This shows that the relevance of search results is highly subjective and can change throughout a search session.

6.1.5 Claim V

Claim V states that user performance improves if IR systems retrieve broad results for exploratory tasks and narrower sets for lookup. Studies IV and V together confirmed the fifth key claim made in this thesis. It is possible to explain why the adaptation of exploration rate truly improves user performance by considering the AIF. With exploratory tasks, users have difficulty in formulating search queries that properly express their information need, so the search engine cannot trust the queries. Most IR systems ignore this fact and rank the documents that are most relevant for the search query at the top. This creates a gap between IR-system-predicted relevance and the actual user-perceived relevance. Particularly in the case of exploratory search tasks, this gap can become very large when user familiarity with the domain is low. This would result in an ecology wherein the result list contains only a few documents that the user perceives as relevant. When, on the other hand, the IR system returns more varied results, through employing a higher exploration rate, there is a better chance of encompassing more topics that the user would perceive as relevant. This affects the search strategy—with determination that more of the documents are relevant in the latter case.

Chapter 7

Discussion

Exploration is a natural human behavior motivated by our thirst for knowledge. With recent leaps forward in technology, we now have the opportunity to access a vast amount of information much more rapidly via the Web [157]. Although many IR systems and technologies have been designed to improve the retrieval of information, information-seekers still struggle when undertaking exploratory searches [7]. Hence, better search systems are important, to enable faster and effortless exploration of information. Accordingly, I investigated exploratory and lookup search tasks for the purpose of revealing solutions to some of the prominent challenges encountered in information search.

This dissertation has addressed the challenges in information search in relation to three themes: 1) conceptualizing and understanding information search, 2) modeling and predicting information search behaviors, and 3) providing real-time adaptive support for exploratory and lookup search tasks. To conceptualize and understand the search strategies applied with different tasks, I used the adaptive interaction framework [123]. The thesis contributes a model designed to predict dynamic information needs—subjective specificity—in exploration, a classifier built from interaction logs to discriminate between exploratory and lookup tasks, and a reinforcement learning model to predict the adaptive interaction strategies pertinent for exploratory search. Possible approaches to provision of real-time support for both exploratory and lookup tasks were proposed through consideration of these models.

7.1 Summary of the Main Findings

The main findings yielded by this research are best summarized with reference to the five key claims.

According to the first claim, exploratory search strategies emerge as an adaptation to ecology, mechanism, and utility in the AIF. In exploratory search tasks, there is no exact "correct" answer; therefore, the utility involves finding documents with high levels of user-expected information gain. This results in an unpredictable distribution with regard to the ecology. Since this scenario involves a search domain that is less familiar to the information-seeker, several cognitive constraints (such as comprehension time and memory) influence the mechanism alongside common constraints such as reading time, fixation time, and saccade time. When undertaking lookup tasks, on the other hand, the user has a clear target in mind, so the utility would generally involve finding a single document. Ecology follows a predictable pattern indicating that the topmost documents are the most relevant here, and the mechanism involves the common cognitive constraints (reading, fixation, and saccade times). When these differences in ecology, mechanism, and utility are borne in mind, it is possible to explain the differences in rational information-seekers' search strategy between exploratory and lookup tasks. Therefore, I was able to propose an approach for implementing a model to predict the rational exploratory search strategies by training a reinforcement learning agent. I applied formal models of the ecology to generate training data, of utility to define the reward function, and of the mechanism to compute the costs involved in the reward function (or human constraints that contribute negative reward). The model-predicted strategies of an exploratory searcher closely match the strategies applied by an actual user. These are exciting findings that confirm the first claim.

The second claim points to exploratory search as the most challenging and one of the most common search purposes. This claim was arrived at via case-study research that involved observation of information-seekers that was followed by a Web-based survey. The study found five common purposes for which academics initiate search activities: staying up to date on the field, exploring unfamiliar topics, collaborating, reviewing literature, and teaching. Of these purposes, exploring unfamiliar topics turned out to be among the most common search purposes yet one of the most difficult to address. These findings emphasize the importance of designing better tools and techniques to support exploration. Although there has been a large amount of interest in designing search systems, of various types, clear room for improvement remains.

Addressing the third claim—that it is possible to model the dynamic nature of information needs in exploratory search and explain how they lead to dynamic search strategies by considering the AIF—I proposed a model to predict how broad or narrow the search results would be deemed with respect to the actual information need of the user (i.e., the subjective specificity) from observable search behaviors. The model proposed is also sensitive to in-session knowledge gain and existing knowledge or experience possessed by the information-seeker. Furthermore, the AIF theoretically validates the logic behind the model. This finding has valuable implications for addressing one of the greatest challenges in the realm of exploratory search: coping with the dynamic knowledge, goals, and information needs.

One of the biggest hindrances encountered in efforts to advance the design of search systems that adaptively support performing both exploratory and lookup tasks is a lack of empirical evidence on separating between the two task types. The next key claim, that it is possible to distinguish the type of search tasks by training a classifier with user interaction data, stemmed from the first claim, whose theoretical validation entailed using the AIF to show how different search strategies can emerge from certain task characteristics. The results of a study conducted to validate the classifier empirically demonstrate how the actual search task can be ascertained while the user is still interacting with the first search-engine result page. This finding provides insight into how one can design adaptive information search systems.

The final phase of the research involved designing an adaptive information search system that provides real-time support for both exploratory and lookup tasks. This system was developed through synthesis of the models developed earlier in the doctoral project. The results of a study that validated this search system allow me to assert that user performance in information search can be improved via adapting of information retrieval algorithms to retrieve broader results for exploratory tasks and narrower results for lookup tasks. The findings from this research allow supporting both task types surreptitiously without departing from the list-based search interfaces familiar to users.

7.2 Implications of the Research

In line with earlier research into information search behaviors with different task types, my findings show that users adapt different information search strategies in keeping with the task type. Yet the findings go further, complementing prior work by revealing the factors that could rationally explain such adaptive search strategies. The adaptive interaction framework that I used for this purpose has been applied for the menu-search context [41], but this is the first time such a framework has been applied to information search tasks to provide thorough explanation as to how and why search behaviors are adaptive. Unlike other models of information search behaviors, which predict only the task completion time [17,124], the computational model created via the AIF is capable of predicting other process variables—clicks and gaze distribution—in addition to task completion time. The proposed reinforcement learning model has important implications for systems that provide support for exploration. The model allows information search systems to make empirical estimates of how new features are going to affect the search behavior. For example, if a search system introduces a feature that highlights relevant keywords, it is possible to interpret how this feature would affect the utility, mechanism, and ecology in line with the AIF. It would then allow us to predict how this feature would change the search strategy.

Chapter 5 presented a model that assesses how broad or narrow the search results are with respect to the actual information need of the user. This model can be applied to exploratory search tasks to provide implicit relevance feedback. The classification results show that our model indeed captures valuable information about the subjective selectivity of results and can be applied in an actual search system. There are several important implications of this model that can be exploited in the design of adaptive search systems. It has potential applications in systems that support exploratory search by making query suggestions [59], organizing information by facet [170], directing search by visualizing relevant keywords [64], and presenting summaries of results in various ways [95]. These systems could use the subjective specificity model to predict whether potential returned results are too broad or too narrow for the user's actual information need. Furthermore, this model could be used as a substitute for relevance feedback techniques that put the user through tedious feedback loops.

The classification study that identified behavioral indicators that separate exploratory tasks from lookup tasks (Section 5.2) also has important implications for the design of adaptive search systems. The classification analysis indicated that the outcome can be exploited and could be used for customizing and adapting IR systems. I posit that several aspects of IR systems can be tailored to the predicted task type: interface design, retrieval algorithm design, and user model design. These aspects are thoroughly discussed in Publication III. Unlike in exploratory searches, in lookup searches users examine fever result items and prefer a shorter informative summary (i.e., snippet) for each result [47]. Since the task type has a significant influence on search behavior, implicit relevance feedback techniques could benefit greatly from classifying the task by type early in the interaction. In a nutshell, the classifier I have proposed allows search systems to adjust the number of result items shown per SERP dynamically, with more result items for exploratory tasks; to alter the snippet length so as to provide longer summaries in cases of exploratory tasks; and to derive implicit relevance feedback more accurately.

The adaptive search system proposed in Chapter 6 provides evidence for the practical applicability of the aforementioned models in real-world search systems. A similar approach could be followed for making adaptations to the search interface and to visible features of the IR system.

7.3 Limitations of the Work

All of the user studies were conducted in laboratory settings. All methods have their drawbacks, and controlled laboratory studies are no exception. Users in the lab are seldom truly motivated to perform the tasks. However, the alternative approach is to collect data from search-engine logs, which would provide little information on the actual task that was performed or about user satisfaction [132], let alone about the task success [14]. Another reason for conducting laboratory studies is that they allow us to control other factors that could affect search behavior. This is vital because search strategies are influenced by many additional factors, such as domain knowledge, search expertise, and task difficulty [99]. I included realistic search tasks and follow-up interviews in all of the studies in order to obtain feedback on actual user perceptions. To compensate for the lack of user motivation to perform externally set tasks, the conditions in the exploratory search studies reported upon in publications II and V allowed the users to explore topics of their interest. Although the models and adaptive search engine developed in this research could be validated by methods utilized in the experiments done in the thesis project, field studies or studies in natural settings are required for addressing the external validity [120].

In all of the studies, the setting for the search tasks was an academic information search context. The primary reason for this selection is that the main goal for exploratory search is to obtain new knowledge, which is particularly important in an academic milieu [165]. In light of this, all the participants in the user studies were researchers. However, validation of the models beyond the academic context necessitates conducting studies in other domains too, with participants who have different backgrounds.

A pool of academic articles from arXiv was the main data source used by the information search system. This corpus was chosen because arXiv is one of the most popular open-access digital libraries in the physics, mathematics, and computer science domains and is a free digital library. Another reason for selection of a set of academic articles is that, again, our studies were conducted in the academic context. Although the modeling approaches and adaptive information search designs proposed in this thesis are suitable for other search tasks as well, we have validated them only with an academic dataset. Generalized studies and information search systems designed with other datasets are necessary for demonstrating the generalizability of the findings.

The classification exercise conducted in the subjective specificity research (Study II) suggested that the model's accuracy improves when the user examines more items on the SERPs (around 33 items). However, larger quantities of training data are required for building a more advanced classifier, one that can make reliable calls while the user is examining the first 10 items on the SERP.

The prototype system developed to provide adaptive support for exploratory and lookup search tasks require the user to provide relevance feedback on the returned documents. However, prior studies show that due to the high cognitive load of providing feedback, such techniques are not in use [88]. Although the relevance feedback is not used in the classifier, the reinforcement learning algorithm requires some sort of feedback to rate the documents. Hence, implicit techniques to acquire user feedback is needed to improve the usability of this system.

The parameters to the reward function in the reinforcement learning model based on the AIF were set by trial and error. Although this approach allowed me to build an accurate model of exploratory search, the accuracy could be further improved by means of parameter optimization techniques. The associated research could benefit also from having more data as input.

7.4 Directions for the Future

In this thesis, the adaptive information search system was assessed through assignment of exploratory and lookup tasks to users in a controlled laboratory setting. Validating the practical applicability of this adaptive search system in a naturalistic setting requires launching the system for use by the public. Controlling various background factors, among them domain knowledge, search expertise, and the difficulty of the tasks, would have demanded overly complicated experiments and extensive additional analysis in order to guarantee realistic tasks and appropriate interpretation of the results. With future work, I intend to investigate how the system developed works in the wild. This represents a complex undertaking: ascertaining the ground truth with respect to the actual search goal that motivated the search is difficult, not least because users must explicitly specify the search task at the outset. On many occasions, the information-seeker has doubts surrounding the type of search task. This requires additional research into methods for uncovering the actual motivation behind the search.

Extension of this line of research beyond the academic search context is another element that calls for further study. Exploratory search occurs in general Web search conditions involving conceptually broad search topics such as a first-time property-buyer exploring the domain [15]. Another typical situations might involve a tourist exploring a new town—a scenario that, while not directly related to search, does involve a form of exploration. One interesting avenue for future investigation would be to generalize the models proposed in this thesis to support exploration in different contexts. Although the application of the proposed approach to general Web search is rather straightforward (since there are no fundamental differences between academic search and general Web search), application to physical exploration would require further investigation. Models such as information foraging theory provide motivation for this line of inquiry by illustrating how a theory derived from the behavior of animals foraging for food can be applied to the completely different context of information search [125]. I plan to examine the possibilities for extending this research beyond the information search context.

In this thesis, I have presented a model to predict the subjective specificity of search results while the user is actively engaged in search. Though several valuable implications of this model have been proposed, it has not yet been integrated into the adaptive search system that I proposed in the course of this research. That is mainly because the adaptation of either the user interface or the retrieval system would require more thorough analysis of user requirements. This step necessitates more careful investigation of how to respond to the subjective specificity and dynamic changes in user knowledge in exploration while simultaneously avoiding over-personalization of the results, since that could lead to context traps [88]. Exploring and critically evaluating the opportunities for dynamic adaptation of IR systems in line with evolving information needs and user knowledge could prove to be a highly fruitful endeavor.

Another important open challenge is how to make the exploration–exploitation tradeoff discussed in Chapter 6 transparent to the user. The approach followed in the doctoral research involved unobtrusively adapting the retrieval algorithm without altering the familiar list-based interface. However, adapting the interface merits more careful investigation in relation to how best to show the system-predicted task type to the user and allow the user to correct the system's prediction. Yet it is also important to avoid overloading the user, since providing too much information could create a distraction [73].

7.5 Conclusion

Exploratory and lookup searches are the two common categories of information search tasks. Although IR systems and technologies have vastly improved over the last decade, information-seekers still need more support in performing exploratory search tasks. There is a need to develop a generalizable conceptualization of information search that helps to distinguish exploratory and lookup searches. Such a conceptualization would enable the development of adaptive IR systems to support both kinds of searches.

In this dissertation, I propose a conceptualization of information search based on an existing framework, called Adaptive Interaction Framework (AIF). Through this conceptualization I explain that the characteristics of exploratory information search strategies are a result of rational choices users make to maximize their information gain (or utility) in a given ecological structure with cognitive and perceptual limits. This conceptualization enables us to explain why exploratory search is challenging and build predictive models.

This research contributes three predictive models of information search behaviors. The first model predicts the dynamic parameters in exploration from eye gaze movement and click interactions. This model is also capable of detecting prior user experience in search, and changes in user knowledge over the course of a search session. The second model is a classifier that distinguishes exploratory and lookup search tasks from implicit user behaviors. The third model is a computationally rational model of adaptive exploratory search behaviors that implements reinforcement learning algorithm to predict optimal information search strategies. These models make important contributions to the development of adaptive IR systems.

The dissertation demonstrates an approach to provide real-time support for both exploratory and lookup search tasks with the predictive models developed in this research. The prototype system shows how to distinguish the search tasks while the user is still searching and dynamically tune the IR system to better match the user needs in each search task.

In conclusion, this thesis has chronicled the gradual development of an approach to providing better support for information search, with important implications. The most important outcomes of this work are a better understanding of exploratory search through the lens of the AIF, designing of models based on this understanding, and an adaptive search system that provides real-time support for information search.

References

[1] Iris Adä and Michael R Berthold. Eve: a framework for event detection. *Evolving systems*, 4(1):61–70, 2013.

[2] Denise E Agosto. Bounded rationality and satisficing in young people's web-based decision making. *Journal of the American Society for Information Science and Technology (JASIST)*, 53(1):16–27, 2002.

[3] Antti Ajanki and Samuel Kaski. Probabilistic proactive timeline browser. In *Artificial Neural Networks and Machine Learning (ICANN)*, pages 357–364. Springer, 2011.

[4] Anwar Alhenshiri, Carolyn Watters, Michael Shepherd, and Jack Duffy. Building support for web information gathering tasks. In *System Science (HICSS)*, pages 1687–1696. IEEE, 2012.

[5] John R Anderson. Is human cognition adaptive? *Behavioral and Brain Sciences*, 14(03):471–485, 1991.

[6] Salvatore Andolina, Khalil Klouche, Diogo Cabral, Tuukka Ruotsalo, and Giulio Jacucci. Inspirationwall: Supporting idea generation through automatic information exploration. In *SIGCHI Conference on Creativity and Cognition (C & C)*, pages 103–106. ACM, 2015.

[7] K. Athukorala, E. Hoggan, A. Lehtiö, T. Ruotsalo, and G. Jacucci. Information-seeking behaviors of computer scientists: Challenges for electronic literature search tools. In *Annual Meeting of the Association for Information Science and Technology (ASIST)*. Wiley Online Library, 2012.

[8] K. Athukorala, A. Medlar, K. Ilves, and D. Głowacka. Balancing exploration and exploitation: Empirical parameterization of exploratory search systems. In *Conference on Information and Knowledge Management (CIKM)*. ACM, 2015.

[9] K. Athukorala, A. Medlar, A. Oulasvirta, G. Jacucci, and D. Głowacka. Beyond relevance : Adapting exploration / exploitation in information retrieval. In *Conference on Intelligent User Interfaces (IUI)*. ACM, 2016.

[10] Kumaripaba Athukorala, Dorota Głowacka, Antti Oulasvirta, Jilles Vreeken, and Giulio Jacucci. Is exploratory search different? a comparison of information search behavior for exploratory and lookup tasks. *Journal of the American Society for Information Science and Technology (JASIST)*, 2015.

[11] Kumaripaba Athukorala, Antti Oulasvirta, Dorota Głowacka, Jilles Vreeken, and Giulio Jacucci. Narrow or broad?: Estimating subjective specificity in exploratory search. In *Conference on Information and Knowledge Management (CIKM)*, pages 819–828. ACM, 2014.

[12] P. Auer. Using confidence bounds for exploitation – exploration tradeoffs. *Journal of Machine Learning Research (JMLR)*, 3:397 – 422, 2002.

[13] A. Aula, R. M. Khan, and Z. Guan. How does search behavior change as search becomes more difficult? In *SIGCHI Conference on Human Factors in Computing Systems (CHI)*, pages 35–45. ACM, 2010.

[14] A. Aula and K. Nordhausen. Modeling successful performance in web searching. *Journal of the American Society for Information Science and Technology (JASIST)*, 57(12):1678–1693, 2006.

[15] Anne Aula and Daniel M Russell. Complex and exploratory web search. In *Information Seeking Support Systems Workshop (ISSS)*, 2008.

[16] Leif Azzopardi. The economics in interactive information retrieval. In *SIGIR conference on Research and development in Information Retrieval (SIGIR)*, pages 15–24. ACM, 2011.

[17] Leif Azzopardi. Modelling interaction with economic models of search. In *SIGIR conference on Research and development in information retrieval (SIGIR)*, pages 3–12. ACM, 2014.

[18] M. Baldonado, Q Wang, and T. Winograd. Sensemaker: an information-exploration interface supporting the contextual evolution of a user's interests. In *SIGCHI conference on Human factors in computing systems (CHI)*, pages 11–18. ACM, 1997.

[19] Feza Baskaya, Heikki Keskustalo, and Kalervo Järvelin. Modeling behavioral factors in interactive information retrieval. In *Conference on information and knowledge management (CIKM)*, pages 2297–2302. ACM, 2013.

[20] M. J. Bates. The design of browsing and berrypicking techniques for the online search interface. *Online Information Review*, 13(5):407–424, 1989.

[21] Nicholas J Belkin. Some (what) grand challenges for information retrieval. In *SIGIR Forum*, volume 42, pages 47–54. ACM, 2008.

[22] Nicholas J. Belkin, Colleen Cool, Diane Kelly, S-J Lin, SY Park, J Perez-Carballo, and C Sikora. Iterative exploration, design and evaluation of support for query reformulation in interactive information retrieval. *Information Processing and Management*, 37(3):403–434, 2001.

[23] David J Bell and Ian Ruthven. Searcher's assessments of task complexity for web searching. In *European Conference on Information Retrieval (ECIR)*, pages 57–71. Springer, 2004.

[24] Albert Bifet and Eibe Frank. Sentiment knowledge discovery in twitter streaming data. In *Discovery Science*, 2010.

[25] Albert Bifet and Ricard Gavalda. Learning from time-changing data with adaptive windowing. In *SIAM Conference on Data Mining (SDM)*, 2007.

[26] David C Blair and Melvin E Maron. An evaluation of retrieval effectiveness for a full-text document-retrieval system. *Communications of the ACM*, 28(3):289–299, 1985.

[27] Pia Borlund and Peter Ingwersen. The development of a method for the evaluation of interactive information retrieval systems. *Journal of documentation*, 53(3):225–250, 1997.

[28] Abdelhamid Bouchachia. Incremental learning with multi-level adaptation. *Neurocomputing*, 74(11):1785–1799, 2011.

[29] S. Brand-Gruwel, I. Wopereis, and Y. Vermetten. Information problem solving by experts and novices: Analysis of a complex cognitive skill. *Computers in Human Behavior*, 21(3):487–508, 2005.

[30] EMC Digital Universe Info Brief. The digital universe of opportunities, 2014.

[31] Andrei Broder. A taxonomy of web search. In *SIGIR Forum*, volume 36, pages 3–10. ACM, 2002.

[32] Katriina Byström. Information and information sources in tasks of varying complexity. *Journal of the American Society for Information Science and Technology (JASIST)*, 53(7):581–591, 2002.

[33] Katriina Byström and Kalervo Järvelin. Task complexity affects information seeking and use. *Information processing and management*, 31(2):191–213, 1995.

[34] Donald J Campbell. Task complexity: A review and analysis. *Academy of management review*, 13(1):40–52, 1988.

[35] Robert Capra, Gary Marchionini, Jung Sun Oh, Fred Stutzman, and Yan Zhang. Effects of structure and interaction style on distinct search tasks. In *ACM/IEEE-CS joint conference on Digital libraries (JCDL)*, pages 442–451. ACM, 2007.

[36] José M Carmona-Cejudo, Manuel Baena-García, José del Campo-Ávila, Rafael Morales-Bueno, and Albert Bifet. Gnusmail: Open framework for on-line email classification. In *Frontiers in Artificial Intelligence and Applications*, 2011.

[37] Claudio Carpineto, Stanislaw Osiński, Giovanni Romano, and Dawid Weiss. A survey of web clustering engines. *ACM Computing Surveys (CSUR)*, 41(3):17, 2009.

[38] Nick Chater and Mike Oaksford. Ten years of the rational analysis of cognition. *Trends in Cognitive Sciences*, 3(2):57–65, 1999.

[39] Nick Chater and Mike Oaksford. The rational analysis of mind and behavior. *Journal for Epistemology, Methodology and Philosophy of Science (Synthese)*, 122(1-2):93–131, 2000.

[40] D. H. Chau, A. Kittur, J. I. Hong, and C. Faloutsos. Apolo: making sense of large network data by combining rich user interaction and machine learning. In *SIGCHI conference on Human factors in computing systems (CHI)*, pages 167–176, 2011.

[41] Xiuli Chen, Gilles Bailly, Duncan P Brumby, Antti Oulasvirta, and Andrew Howes. The emergence of interactive behaviour: A model of

rational menu search. In *SIGCHI Conference on Human Factors in Computing Systems (CHI)*, pages 4217–4226. ACM, 2015.

[42] Sudatta Chowdhury, Forbes Gibb, and Monica Landoni. Uncertainty in information seeking and retrieval: A study in an academic environment. *Information Processing and Management*, 47(2):157–175, 2011.

[43] Charles LA Clarke, Maheedhar Kolla, Gordon V Cormack, Olga Vechtomova, Azin Ashkan, Stefan Büttcher, and Ian MacKinnon. Novelty and diversity in information retrieval evaluation. In *SIGIR conference on Research and development in information retrieval (SIGIR)*, pages 659–666. ACM, 2008.

[44] M. J. Cole, J. Gwizdka, C. Liu, N. J. Belkin, and X. Zhang. Inferring user knowledge level from eye movement patterns. *Information Processing and Management*, 2012.

[45] Michael D Cooper. A cost model for evaluating information retrieval systems. *Journal of the American Society for Information Science and Technology (JASIST)*, 23(5):306–312, 1972.

[46] Julie Courraud. Zotero: A free and open-source reference manager. *Medical Writing*, 23(1):46–48, 2014.

[47] E. Cutrell and Z. Guan. What are you looking for?: An eye-tracking study of information usage in web search. In *SIGCHI Conference on Human Factors in Computing Systems (CHI)*, pages 407–416, 2007.

[48] D. R. Cutting, D. R Karger, J. O. Pedersen, and J. W. Tukey. Scatter/gather: A cluster-based approach to browsing large document collections. In *SIGIR conference on Research and development in Information Retrieval (SIGIR)*, pages 318–329, 1992.

[49] Douglass R Cutting, David R Karger, Jan O Pedersen, and John W Tukey. Scatter/gather: A cluster-based approach to browsing large document collections. In *SIGIR conference on Research and development in Information Retrieval (SIGIR)*, pages 318–329. ACM, 1992.

[50] A. M. Diriye. *Search interfaces for known-item and exploratory search tasks*. PhD thesis, UCL (University College London), 2012.

[51] Abdigani Diriye, Ann Blandford, and Anastasios Tombros. Exploring the impact of search interface features on search tasks. In *Research and Advanced Technology for Digital Libraries*, pages 184–195. Springer, 2010.

[52] Abdigani Diriye, Ann Blandford, and Anastasios Tombros. When is system support effective? In *Symposium on Information interaction in context (IIiX)*, pages 55–64. ACM, 2010.

[53] Abdigani Diriye, Max L Wilson, Ann Blandford, and Anastasios Tombros. Revisiting exploratory search from the hci perspective. In *Workshop on Human-Computer Interaction and Information Retrieval (HCIR)*, page 99. Citeseer, 2010.

[54] Jia Tina Du and Nina Evans. Academic users' information searching on research topics: Characteristics of research tasks and search strategies. *The journal of academic librarianship*, 37(4):299–306, 2011.

[55] Carsten Eickhoff, Jaime Teevan, Ryen White, and Susan Dumais. Lessons from the journey: a query log analysis of within-session learning. In *Conference on Web search and data mining (WSDM)*, pages 223–232. ACM, 2014.

[56] EMC. Yandex just passed bing to become 4th largest global search engine, 2013.

[57] Ronald Aylmer Fisher. *The design of experiments*, volume 12. Oliver and Boyd Edinburgh, 1960.

[58] P. M. Fitts and J. R. Peterson. Information capacity of discrete motor responses. *Journal of experimental psychology*, 67(2):103, 1964.

[59] W.-T. Fu, T. G. Kannampallil, and R. Kang. Facilitating exploratory search by model-based navigational cues. In *Conference on intelligent user interfaces (IUI)*, 2010.

[60] W.-T. Fu and P. Pirolli. Snif-act: A cognitive model of user navigation on the world wide web. *Human-Computer Interaction*, 22(4):355–412, 2007.

[61] Norbert Fuhr. A probability ranking principle for interactive information retrieval. *Information Retrieval*, 11(3):251–265, 2008.

[62] João Gama, Indrė Žliobaitė, Albert Bifet, Mykola Pechenizkiy, and Abdelhamid Bouchachia. A survey on concept drift adaptation. *ACM Computing Surveys (CSUR)*, 46(4):44:1–44:37, 2014.

[63] Samuel J Gershman, Eric J Horvitz, and Joshua B Tenenbaum. Computational rationality: A converging paradigm for intelligence in brains, minds, and machines. *Science*, 349(6245):273–278, 2015.

[64] D. Głowacka, T. Ruotsalo, K. Konyushkova, K. Athukorala, S. Kaski, and G. Jacucci. Directing exploratory search: Reinforcement learning from user interactions with keywords. In *Conference on Intelligent User Interfaces (IUI)*, 2013.

[65] R. L. Goldstone and B. C. Ashpole. Human foraging behavior in a virtual environment. *Psychonomic Bulletin and Review*, 11(3):508–514, 2004.

[66] Gene Golovchinsky. What the query told the link: the integration of hypertext and information retrieval. In *Conference on Hypertext*, pages 67–74. ACM, 1997.

[67] Gene Golovchinsky, Pernilla Qvarfordt, and Jeremy Pickens. Collaborative information seeking. In *NSF Workshop on Information Seeking Support Systems*, number 3, pages 47–51. IEEE, 2009.

[68] João Bártolo Gomes, Ernestina Menasalvas, and Pedro AC Sousa. Learning recurring concepts from data streams with a context-aware ensemble. In *Symposium on applied computing*, 2011.

[69] L. A Granka, T. Joachims, and G. Gay. Eye-tracking analysis of user behavior in www search. In *SIGIR conference on Research and development in Information Retrieval (SIGIR)*, 2004.

[70] Farrokh Habibzadeh and Mahboobeh Yadollahie. Are shorter article titles more attractive for citations? crosssectional study of 22 scientific journals. *Croatian medical journal*, 51(2):165–170, 2010.

[71] A. Hassan and R. W. White. Task tours: helping users tackle complex search tasks. In *Conference on Information and Knowledge Management (CIKM)*, 2012.

[72] A. Hassan, R. W. White, S. T. Dumais, and Y. Wang. Struggling or exploring?: disambiguating long search sessions. In *Conference on Web Search and Data Mining (WSDM)*, pages 53–62, 2014.

[73] M. Hearst. *Search user interfaces*. Cambridge University Press, 2009.

[74] Marti A Hearst. Clustering versus faceted categories for information exploration. *Communications of the ACM*, 49(4):59–61, 2006.

[75] Bradley M Hemminger, Dihui Lu, KTL Vaughan, and Stephanie J Adams. Information seeking behavior of academic scientists. *Journal of the American Society for Information Science and Technology (JASIST)*, 58(14):2205–2225, 2007.

[76] William Hersh, Jeffrey Pentecost, and David Hickam. A task-oriented approach to information retrieval evaluation. *Journal of American Society for Information Science (JASIS)*, 47(1):50–56, 1996.

[77] Christoph Hölscher and Gerhard Strube. Web search behavior of internet experts and newbies. *Computer Networks*, 33(1):337–346, 2000.

[78] Peter Ingwersen and Kalervo Järvelin. *The turn: Integration of information seeking and retrieval in context*, volume 18. Springer Science & Business Media, 2006.

[79] B. J. Jansen, D. L. Booth, and A. Spink. Determining the informational, navigational, and transactional intent of web queries. *Information Processing and Management*, 44(3):1251–1266, 2008.

[80] Bernard J Jansen and Udo Pooch. A review of web searching studies and a framework for future research. *Journal of the American Society for Information Science and Technology (JASIST)*, 52(3):235–246, 2001.

[81] Kalervo Järvelin and TD Wilson. On conceptual models for information seeking and retrieval research. *Information Research*, 9(1):9–1, 2003.

[82] Christine Jenkins, Cynthia L Corritore, and Susan Wiedenbeck. Patterns of information seeking on the web: A qualitative study of domain expertise and web expertise. *IT and society*, 1(3):64–89, 2003.

[83] T. Joachims, L. Granka, B. Pan, H. Hembrooke, and G. Gay. Accurately interpreting clickthrough data as implicit feedback. In *SIGIR conference on Research and development in Information Retrieval (SIGIR)*, pages 154–161, 2005.

[84] Rosie Jones, Benjamin Rey, Omid Madani, and Wiley Greiner. Generating query substitutions. In *Conference on World Wide Web (WWW)*, pages 387–396. ACM, 2006.

[85] M. Käki. Findex: search result categories help users when document ranking fails. In *SIGCHI conference on Human factors in computing systems (CHI)*, pages 131–140. ACM, 2005.

[86] Maryam Kamvar, Melanie Kellar, Rajan Patel, and Ya Xu. Computers and iphones and mobile phones, oh my!: a logs-based comparison of search users on different devices. In *Conference on World Wide Web (WWW)*, pages 801–810. ACM, 2009.

[87] Richard M Keller, Shawn R Wolfe, James R Chen, Joshua L Rabinowitz, and Nathalie Mathe. A bookmarking service for organizing and sharing urls. *Computer Networks and ISDN Systems*, 29(8):1103–1114, 1997.

[88] D. Kelly and X. Fu. Elicitation of term relevance feedback: an investigation of term source and context. In *SIGIR conference on Research and development in information retrieval (SIGIR)*, pages 453–460. ACM, 2006.

[89] J. Kim. Describing and predicting information-seeking behavior on the web. *Journal of the American Society for Information Science and Technology (JASIST)*, 60(4):679–693, 2009.

[90] Kyung-Sun Kim. Information-seeking on the web: Effects of user and task variables. *Library and Information Science Research*, 23(3):233–255, 2001.

[91] Kyung-Sun Kim and Bryce Allen. Cognitive and task influences on web searching behavior. *Journal of the American Society for Information Science and Technology (JASIST)*, 53(2):109–119, 2002.

[92] Walter Kintsch. *Comprehension: A paradigm for cognition.* Cambridge university press, 1998.

[93] Khalil Klouche, Tuukka Ruotsalo, Diogo Cabral, Salvatore Andolina, Andrea Bellucci, and Giulio Jacucci. Designing for exploratory search on touch devices. In *SIGCHI Conference on Human Factors in Computing Systems (CHI)*, pages 4189–4198. ACM, 2015.

[94] Carol C Kuhlthau. Inside the search process: Information seeking from the user's perspective. *Journal of the American society for information science (JASIS)*, 42(5):361–371, 1991.

[95] B. Kules, M. Wilson, M. C. Schraefel, B. Shneiderman, et al, and et al. From keyword search to exploration: How result visualization aids discovery on the web. Technical Report HCIL-2008-06, University of Maryland, 2008.

[96] Bill Kules and Robert Capra. Creating exploratory tasks for a faceted search interface. *Human-Computer Interaction and Information Retrieval (HCIR)*, pages 18–21, 2008.

[97] Bill Kules and Ben Shneiderman. Users can change their web search tactics: Design guidelines for categorized overviews. *Information Processing and Management*, 44(2):463–484, 2008.

[98] L. Li, W. Chu, J. Langford, and R.E. Schapire. A contextual-bandit approach to personalized news article recommendation. In *conference on World Wide Web (WWW)*, 2010.

[99] Yuelin Li and Nicholas J Belkin. A faceted approach to conceptualizing tasks in information seeking. *Information Processing and Management*, 44(6):1822–1837, 2008.

[100] Yuelin Li and Nicholas J Belkin. An exploration of the relationships between work task and interactive information search behavior. *Journal of the American Society for Information Science and Technology (JASIST)*, 61(9):1771–1789, 2010.

[101] Siân E Lindley, Sam Meek, Abigail Sellen, and Richard Harper. It's simply integral to what i do: enquiries into how the web is weaved into everyday life. In *conference on World Wide Web (WWW)*, pages 1067–1076. ACM, 2012.

[102] Patrick Lindstrom, Sarah Jane Delany, and Brian Mac Namee. Handling concept drift in text data stream constrained by high labelling cost. In *Florida Artificial Intelligence Research Society Conference (FLAIRS)*, 2010.

[103] Lars Lischke, Jürgen Grüninger, Khalil Klouche, Albrecht Schmidt, Philipp Slusallek, and Giulio Jacucci. Interaction techniques for wall-sized screens. In *conference on Interactive Tabletops and Surfaces (ITS)*, pages 501–504. ACM, 2015.

[104] Jingjing Liu, Michael J Cole, Chang Liu, Ralf Bierig, Jacek Gwizdka, Nicholas J Belkin, Jun Zhang, and Xiangmin Zhang. Search behaviors in different task types. In *Joint Conference on Digital Libraries (JCDL)*, pages 69–78. ACM, 2010.

[105] Jingjing Liu, Chang Liu, and Nicholas J Belkin. Predicting information searchers' topic knowledge at different search stages. *Journal of the Association for Information Science and Technology (JASIST)*, 2016.

[106] Jingjing Liu, Chang Liu, Jacek Gwizdka, and Nicholas J Belkin. Can search systems detect users' task difficulty?: Some behavioral signals. In *SIGIR conference on Research and development in Information Retrieval (SIGIR)*, pages 845–846. ACM, 2010c.

[107] G. Marchionini. Exploratory search: from finding to understanding. *Communications of the ACM*, 49(4):41–46, 2006.

[108] Gary Marchionini. Information-seeking strategies of novices using a full-text electronic encyclopedia. *SIGIR conference on Research and development in Information Retrieval (SIGIR)*, 40(1):54–66, 1989.

[109] Gary Marchionini, Sandra Dwiggins, Andrew Katz, and Xia Lin. Information seeking in full-text end-user-oriented search systems: The roles of domain and search expertise. *Library and information science research*, 15(1):35–69, 1993.

[110] J. Matejka, T. Grossman, and G. Fitzmaurice. Citeology: visualizing paper genealogy. In *CHI Extended Abstracts*, pages 181–190. ACM, 2012.

[111] E Mcgrath. Methodology matters: Doing research in the behavioral and social sciences. In *Readings in Human-Computer Interaction: Toward the Year 2000*. Citeseer, 1995.

[112] David Millen, Jonathan Feinberg, and Bernard Kerr. Social bookmarking in the enterprise, 2006.

[113] Meredith Ringel Morris. Interfaces for collaborative exploratory web search: Motivations and directions for multi-user design. In *CHI Workshop on Exploratory Search and HCI: Designing and Evaluating Interfaces to Support Exploratory Search Interaction*, pages 9–12, 2007.

[114] Meredith Ringel Morris and Eric Horvitz. Searchtogether: an interface for collaborative web search. In *Symposium on User Interface Software and Technology*, pages 3–12. ACM, 2007.

[115] Meredith Ringel Morris and Jaime Teevan. Collaborative web search: Who, what, where, when, and why. *Synthesis Lectures on Information Concepts, Retrieval, and Services*, 1(1):1–99, 2009.

[116] R. Navarro-Prieto, M. Scaife, and Y. Rogers. Cognitive strategies in web searching. In *Human Fact. and the Web*, pages 43–56, 1999.

[117] Xi Niu and Bradley M Hemminger. A study of factors that affect the information-seeking behavior of academic scientists. *Journal of the American Society for Information Science and Technology (JASIST)*, 63(2):336–353, 2012.

[118] Yuan Niu and Susan Winter. The effect of task type and information format on web searching performance. *International Conference on Information Systems (ICIS)*, page 52, 2006.

[119] Mike Oaksford and Nick Chater. A rational analysis of the selection task as optimal data selection. *Psychological Review*, 101(4):608, 1994.

[120] Antti Oulasvirta. Rethinking experimental designs for field evaluations. *IEEE Pervasive Computing*, (4):60–67, 2012.

[121] R. A. Palmquist and K.-S. Kim. Cognitive style and on-line database search experience as predictors of web search performance. *Journal of the American Society for Information Science and Technology (JASIST)*, 51(6):558–566, 2000.

[122] Emily S Patterson, Emilie M Roth, and David D Woods. Predicting vulnerabilities in computer-supported inferential analysis under data overload. *Cognition, Technology and Work*, 3(4):224–237, 2001.

[123] Stephen J Payne and Andrew Howes. Adaptive interaction: a utility maximization approach to understanding human interaction with technology. *Synthesis Lectures on Human-Centered Informatics*, 6(1):1–111, 2013.

[124] P. Pirolli and S. Card. Information foraging. *Psychological review*, 106(4):643, 1999.

[125] Peter Pirolli. *Information foraging theory: Adaptive interaction with information.* Oxford University Press, 2007.

[126] Wanda Pratt, Marti A Hearst, and Lawrence M Fagan. A knowledge-based approach to organizing retrieved documents. In *American Association for Artificial Intelligence (AAAI)*, pages 80–85, 1999.

[127] Liwen Qiu. Analytical searching vs. browsing in hypertext information retrieval systems. *Canadian Journal of Information and Library Science*, 18(4):1–13, 1993.

[128] Yan Qu and George W Furnas. Model-driven formative evaluation of exploratory search: A study under a sensemaking framework. *Information Processing and Management*, 44(2):534–555, 2008.

[129] F. Radlinski, R. Kleinberg, and T. Joachims. Learning diverse rankings with multi-armed bandits. In *International Conference on Machine Learning (ICML)*, pages 784–791, 2008.

[130] Keith Rayner. Eye movements in reading and information processing: 20 years of research. *Psychological bulletin*, 124(3):372, 1998.

[131] Stephen E Robertson. The probability ranking principle in ir. *Journal of documentation*, 33(4):294–304, 1977.

[132] Daniel E Rose and Danny Levinson. Understanding user goals in web search. In *conference on World Wide Web (WWW)*, pages 13–19. ACM, 2004.

[133] Tuukka Ruotsalo, Kumaripaba Athukorala, Dorota Głowacka, Ksenia Konyushkova, Antti Oulasvirta, Samuli Kaipiainen, Samuel Kaski, and Giulio Jacucci. Supporting exploratory search tasks with interactive user modeling. In *Annual Meeting of the Association for Information Science and Technology (ASIST)*, number 1, pages 1–10. Wiley Online Library, 2013.

[134] H. Saito and K. Miwa. A cognitive study of information seeking processes in the WWW: The effects of searcher's knowledge and experience. In *Conference on Web Information Systems Engineering (WISE)*, volume 1, pages 321–327. IEEE, 2001.

[135] Gerard Salton and Chris Buckley. Improving retrieval performance by relevance feedback. *Readings in information retrieval*, 24(5):355–363, 1997.

[136] John Schacter, Gregory KWK Chung, and Aimée Dorr. Children's internet searching on complex problems: performance and process analyses. *Journal of the American society for Information Science (JASIST)*, 49(9):840–849, 1998.

[137] Chirag Shah, Chathra Hendahewa, and Roberto González-Ibáñez. Rain or shine? forecasting search process performance in exploratory search tasks. *Journal of the American Society for Information Science and Technology (JASIST)*, 2015.

[138] Herbert A Simon. Invariants of human behavior. *Annual review of psychology*, 41(1):1–20, 1990.

[139] Georg Singer, Ulrich Norbisrath, Eero Vainikko, Hannu Kikkas, and Dirk Lewandowski. Search-logger analyzing exploratory search tasks. In *Symposium on Applied Computing*, pages 751–756. ACM, 2011.

[140] Alan F Smeaton, Hyowon Lee, Colum Foley, Sinéad McGivney, and Cathal Gurrin. Físchlár-diamondtouch: Collaborative video searching on a table. In *Electronic Imaging 2006*, pages 607308–607308. International Society for Optics and Photonics, 2006.

[141] K. Sparck Jones, S. Walker, and S. E. Robertson. A probabilistic model of information retrieval: Development and comparative experiments. *Information Processing and Management*, 36(6):779–840, 2000.

[142] Richard S Sutton and Andrew G Barto. *Reinforcement learning: An introduction*. MIT press, 1998.

[143] Carol Tenopir. Use and users of electronic library resources: An overview and analysis of recent research studies. *Council on Library and Information Resources*, 2003.

[144] A. Thatcher. Web search strategies: The influence of web experience and task type. *Information Processing and Management*, 44(3):1308–1329, 2008.

[145] Yuan-Chi Tseng and Andrew Howes. The adaptation of visual search to utility, ecology and design. *International Journal of Human-Computer Studies*, 80:45–55, 2015.

[146] Daniel Tunkelang. Faceted search. *Synthesis lectures on information concepts, retrieval, and services*, 1(1):1–80, 2009.

[147] P. Vakkari. Task complexity, problem structure and information actions: integrating studies on information seeking and retrieval. *Information processing and management*, 35(6):819–837, 1999.

[148] P. Vakkari. Task-based information searching. *Annual Review of Information Science and Technology*, 37(1):413–464, 2003.

[149] Pertti Vakkari. Exploratory searching as conceptual exploration. *Proceedings of HCIR*, pages 24–27, 2010.

[150] Pertti Vakkari. Searching as learning: A systematization based on literature. *Journal of Information Science*, 42(1):7–18, 2016.

[151] Hal R Varian. Economics and search. In *SIGIR Forum*, volume 33, pages 1–5, 1999.

[152] Hal R Varian and Jack Repcheck. *Intermediate microeconomics: a modern approach*, volume 7. WW Norton New York, 2010.

[153] Nicolas Vibert, Christine Ros, Ludovic Le Bigot, Mélanie Ramond, Jérôme Gatefin, and Jean-François Rouet. Effects of domain knowledge on reference search with the pubmed database: An experimental study. *Journal of the American Society for Information Science and Technology (JASIST)*, 60(7):1423–1447, 2009.

[154] Robert Villa, Iván Cantador, Hideo Joho, and Joemon M Jose. An aspectual interface for supporting complex search tasks. In *SIGIR conference on Research and development in information retrieval (SIGIR)*, pages 379–386. ACM, 2009.

[155] Taowei David Wang, Catherine Plaisant, Alexander J Quinn, Roman Stanchak, Shawn Murphy, and Ben Shneiderman. Aligning temporal data by sentinel events: discovering patterns in electronic health records. In *SIGCHI conference on Human factors in computing systems (CHI)*, pages 457–466. ACM, 2008.

[156] R. W. White, B. Kules, and S. M. Drucker. Supporting exploratory search. *Communications of the ACM*, 49(4):36–39, 2006.

[157] R. W. White and R. A. Roth. Exploratory search: Beyond the query-response paradigm. *Morgan & Claypool Series on Information Concepts, Retrieval, and Services.*, 1(1):1–98, 2009.

[158] Ryen W White. *Interactions with search systems.* Cambridge University Press, 2016.

[159] Ryen W White and Raman Chandrasekar. Exploring the use of labels to shortcut search trails. In *SIGIR conference on Research and development in Information Retrieval (SIGIR)*, pages 811–812. ACM, 2010.

[160] Ryen W White, Joemon M Jose, and Ian Ruthven. Using top-ranking sentences to facilitate effective information access. *Journal of the American Society for Information Science and Technology (JASIST)*, 56(10):1113–1125, 2005.

[161] Ryen W White, Bill Kules, and Ben Bederson. Exploratory search interfaces: categorization, clustering and beyond: report on the xsi 2005 workshop at the human-computer interaction laboratory, university of maryland. In *ACM SIGIR Forum*, volume 39, pages 52–56. ACM, 2005.

[162] Ryen W White and Gary Marchionini. Examining the effectiveness of real-time query expansion. *Information Processing and Management*, 43(3):685–704, 2007.

[163] Ryen W White, Gheorghe Muresan, and Gary Marchionini. Report on acm sigir 2006 workshop on evaluating exploratory search systems. In *SIGIR Forum*, volume 40, pages 52–60. ACM, 2006.

[164] Ryen W White, Hyunyoung Song, and Jay Liu. Concept maps to support oral history search and use. In *ACM/IEEE-CS joint conference on Digital libraries (JCDL)*, pages 192–193. ACM, 2006.

[165] Barbara M Wildemuth and Luanne Freund. Assigning search tasks designed to elicit exploratory search behaviors. In *Human-Computer Interaction and Information Retrieval (HCIR)*, pages 1–10. ACM, 2012.

[166] Max L Wilson. *An analytical inspection framework for evaluating the search tactics and user profiles supported by information seeking interfaces*. PhD thesis, University of Southampton, 2009.

[167] Max L Wilson, Chaoyu Ye, Michael B Twidale, Hannah Grasse, Jacob Rosenthal, and Max McKittrick. Search literacy: learning to search to learn. In *Workshop on Search as Learning (SAL)*, 2016.

[168] Tom D Wilson. Models in information behaviour research. *Journal of documentation*, 55(3):249–270, 1999.

[169] Wan-Ching Wu, Diane Kelly, Ashlee Edwards, and Jaime Arguello. Grannies, tanning beds, tattoos and nascar: Evaluation of search tasks with varying levels of cognitive complexity. In *Information Interaction in Context Symposium*, pages 254–257. ACM, 2012.

[170] K. P. Yee, K. Swearingen, K. Li, and M. Hearst. Faceted metadata for image search and browsing. In *SIGCHI conference on Human factors in computing systems (CHI)*. ACM, 2003.

[171] Yisong Yue and Thorsten Joachims. Interactively optimizing information retrieval systems as a dueling bandits problem. In *International Conference on Machine Learning (ICML)*, 2009.

[172] X Zhang, J Liu, M Cole, and N.J. Belkin. Modeling domain knowledge from searchers' behaviors using multiple regression analysis. *Journal of the American Society for Information Science and Technology (JASIST)*, 66(5):980–1000, 2015.

[173] Xiangmin Zhang, Jingjing Liu, and Michael Cole. Task topic knowledge vs. background domain knowledge: Impact of two types of knowledge on user search performance. In *Advances in Information Systems and Technologies*, pages 179–191. Springer, 2013.

[174] Yuxiang Zhu, David Modjeska, Daniel Wigdor, Shengdong Zhao, et al. Hunter gatherer: interaction support for the creation and management of within-web-page collections. In *conference on World Wide Web (WWW)*, pages 172–181. ACM, 2002.

TIETOJENKÄSITTELYTIETEEN LAITOS
PL 68 (Gustaf Hällströmin katu 2 b)
00014 Helsingin yliopisto

DEPARTMENT OF COMPUTER SCIENCE
P.O. Box 68 (Gustaf Hällströmin katu 2 b)
FI-00014 University of Helsinki, FINLAND

JULKAISUSARJA **A**

SERIES OF PUBLICATIONS **A**

Reports are available on the e-thesis site of the University of Helsinki.

A-2011-1 A. Tripathi: Data Fusion and Matching by Maximizing Statistical Dependencies. 89+109 pp. (Ph.D. Thesis)

A-2011-2 E. Junttila: Patterns in Permuted Binary Matrices. 155 pp. (Ph.D. Thesis)

A-2011-3 P. Hintsanen: Simulation and Graph Mining Tools for Improving Gene Mapping Efficiency. 136 pp. (Ph.D. Thesis)

A-2011-4 M. Ikonen: Lean Thinking in Software Development: Impacts of Kanban on Projects. 104+90 pp. (Ph.D. Thesis)

A-2012-1 P. Parviainen: Algorithms for Exact Structure Discovery in Bayesian Networks. 132 pp. (Ph.D. Thesis)

A-2012-2 J. Wessman: Mixture Model Clustering in the Analysis of Complex Diseases. 118 pp. (Ph.D. Thesis)

A-2012-3 P. Pöyhönen: Access Selection Methods in Cooperative Multi-operator Environments to Improve End-user and Operator Satisfaction. 211 pp. (Ph.D. Thesis)

A-2012-4 S. Ruohomaa: The Effect of Reputation on Trust Decisions in Inter-enterprise Collaborations. 214+44 pp. (Ph.D. Thesis)

A-2012-5 J. Sirén: Compressed Full-Text Indexes for Highly Repetitive Collections. 97+63 pp. (Ph.D. Thesis)

A-2012-6 F. Zhou: Methods for Network Abstraction. 48+71 pp. (Ph.D. Thesis)

A-2012-7 N. Välimäki: Applications of Compressed Data Structures on Sequences and Structured Data. 73+94 pp. (Ph.D. Thesis)

A-2012-8 S. Varjonen: Secure Connectivity With Persistent Identities. 139 pp. (Ph.D. Thesis)

A-2012-9 M. Heinonen: Computational Methods for Small Molecules. 110+68 pp. (Ph.D. Thesis)

A-2013-1 M. Timonen: Term Weighting in Short Documents for Document Categorization, Keyword Extraction and Query Expansion. 53+62 pp. (Ph.D. Thesis)

A-2013-2 H. Wettig: Probabilistic, Information-Theoretic Models for Etymological Alignment. 130+62 pp. (Ph.D. Thesis)

A-2013-3 T. Ruokolainen: A Model-Driven Approach to Service Ecosystem Engineering. 232 pp. (Ph.D. Thesis)

A-2013-4 A. Hyttinen: Discovering Causal Relations in the Presence of Latent Confounders. 107+138 pp. (Ph.D. Thesis)

A-2013-5 S. Eloranta: Dynamic Aspects of Knowledge Bases. 123 pp. (Ph.D. Thesis)

A-2013-6 M. Apiola: Creativity-Supporting Learning Environments: Two Case Studies on Teaching Programming. 62+83 pp. (Ph.D. Thesis)

A-2013-7 T. Polishchuk: Enabling Multipath and Multicast Data Transmission in Legacy and Future Interenet. 72+51 pp. (Ph.D. Thesis)

A-2013-8 P. Luosto: Normalized Maximum Likelihood Methods for Clustering and Density Estimation. 67+67 pp. (Ph.D. Thesis)

A-2013-9 L. Eronen: Computational Methods for Augmenting Association-based Gene Mapping. 84+93 pp. (Ph.D. Thesis)

A-2013-10 D. Entner: Causal Structure Learning and Effect Identification in Linear Non-Gaussian Models and Beyond. 79+113 pp. (Ph.D. Thesis)

A-2013-11 E. Galbrun: Methods for Redescription Mining. 72+77 pp. (Ph.D. Thesis)

A-2013-12 M. Pervilä: Data Center Energy Retrofits. 52+46 pp. (Ph.D. Thesis)

A-2013-13 P. Pohjalainen: Self-Organizing Software Architectures. 114+71 pp. (Ph.D. Thesis)

A-2014-1 J. Korhonen: Graph and Hypergraph Decompositions for Exact Algorithms. 62+66 pp. (Ph.D. Thesis)

A-2014-2 J. Paalasmaa: Monitoring Sleep with Force Sensor Measurement. 59+47 pp. (Ph.D. Thesis)

A-2014-3 L. Langohr: Methods for Finding Interesting Nodes in Weighted Graphs. 70+54 pp. (Ph.D. Thesis)

A-2014-4 S. Bhattacharya: Continuous Context Inference on Mobile Platforms. 94+67 pp. (Ph.D. Thesis)

A-2014-5 E. Lagerspetz: Collaborative Mobile Energy Awareness. 60+46 pp. (Ph.D. Thesis)

A-2015-1 L. Wang: Content, Topology and Cooperation in In-network Caching. 190 pp. (Ph.D. Thesis)

A-2015-2 T. Niinimäki: Approximation Strategies for Structure Learning in Bayesian Networks. 64+93 pp. (Ph.D. Thesis)

A-2015-3 D. Kempa: Efficient Construction of Fundamental Data Structures in Large-Scale Text Indexing. 68+88 pp. (Ph.D. Thesis)

A-2015-4 K. Zhao: Understanding Urban Human Mobility for Network Applications. 62+46 pp. (Ph.D. Thesis)

A-2015-5 A. Laaksonen: Algorithms for Melody Search and Transcription. 36+54 pp. (Ph.D. Thesis)

A-2015-6 Y. Ding: Collaborative Traffic Offloading for Mobile Systems. 223 pp. (Ph.D. Thesis)

A-2015-7 F. Fagerholm: Software Developer Experience: Case Studies in Lean-Agile and Open Source Environments. 118+68 pp. (Ph.D. Thesis)

A-2016-1 T. Ahonen: Cover Song Identification using Compression-based Distance Measures. 122+25 pp. (Ph.D. Thesis)

A-2016-2 O. Gross: World Associations as a Language Model for Generative and Creative Tasks. 60+10+54 pp. (Ph.D. Thesis)

A-2016-3 J. Määttä: Model Selection Methods for Linear Regression and Phylogenetic Reconstruction. 44+73 pp. (Ph.D. Thesis)

A-2016-4 J. Toivanen: Methods and Models in Linguistic and Musical Computational Creativity. 56+8+79 pp. (Ph.D. Thesis)