UNIVERSITÉ CÔTE D'AZUR

ÉCOLE DOCTORALE
SCIENCES ET TECHNOLOGIES DE L'INFORMATION ET DE LA COMMUNICATION

THÈSE DE DOCTORAT

Evaluation des moteurs de recherche exploratoire :

Elaboration d'un corps de méthodes centrées utilisateurs, basées sur une modélisation du processus de recherche exploratoire

Emilie PALAGI

Wimmics (Inria, I3S) et Data Science (EURECOM)

Présentée en vue de l'obtention
du grade de docteur en Informatique
d'Université Côte d'Azur
Dirigée par : Fabien Gandon
Co-encadrée par : Alain Giboin
et Raphaël Troncy
Soutenue le : 23 novembre 2018

Devant le jury, composé de :
Christian Bastien, Professeur, Université de Loraine
Pierre De Loor, Professeur, ENIB
Sylvie Catellin, Maître de conférences, UVSQ
Fabien Gandon, Directeur de recherche, Inria
Alain Giboin, Chargé de recherche, Inria
Raphaël Troncy, Maître de conférences, EURECOM

LABEX UCN User Centric Network @ SOPHIA

Inria inventeurs du monde numérique

Evaluation des moteurs de recherche exploratoire :

Elaboration d'un corps de méthodes centrées utilisateurs, basées sur une modélisation du processus de recherche exploratoire

---

Evaluating exploratory search engines:

Designing a set of user-centered methods based on a modeling of the exploratory search process

Jury :

Directeur de thèse / Thesis supervisor

Fabien Gandon, Directeur de recherche, Inria

Rapporteurs / Reviewers

Christian Bastien, Professeur, Université de Loraine

Pierre De Loor, Professeur, Ecole Nationale d'Ingénieurs de Brest (ENIB)

Examinatrice / Examiner

Sylvie Catellin, Maître de conférences (HDR), Université de Versailles-Saint-Quentin-en-Yvelines (UVSQ)

Invités (co-encadrants) / Guests (co-supervisors)

Alain Giboin, Chargé de recherche, Inria

Raphaël Troncy, Maître de conférences, EURECOM

Résumé

Les moteurs de recherche exploratoire (MRE) sont des logiciels aidant les utilisateurs à explorer un domaine d'intérêt pour y faire des découvertes. Ces moteurs se distinguent en cela des moteurs de recherche classiques tels que Google, Bing ou Yahoo!, lesquels supportent plutôt des recherches ciblées ou recherches de consultation (lookup). Si l'on admet que l'évaluation des MRE vise à vérifier si ces derniers aident effectivement les utilisateurs à réaliser leur tâche d'exploration, on constate que les méthodes existantes d'évaluation de ces systèmes ne permettent pas réellement cette vérification. L'une des raisons à cela est que ces méthodes ne reposent pas sur un modèle approprié de la recherche exploratoire (RE) ou qu'elles restent accrochées à un modèle de la recherche de consultation. L'objectif principal de cette thèse est de proposer aux concepteurs de ces MRE des méthodes d'évaluation centrées utilisateurs reposant sur un modèle du processus de RE. Ainsi, après avoir modélisé le processus de RE, nous proposons deux méthodes d'évaluation qui peuvent être utilisées tout au long du processus de conception. La première méthode, une méthode d'inspection sans utilisateurs, peut être utilisée dès les premières maquettes, et repose sur des heuristiques de RE. Nous avons également proposé des outils facilitant l'utilisation de ces heuristiques : un formulaire en ligne ainsi qu'une extension Google Chrome appelée CheXplore. La seconde méthode, avec utilisateurs, peut être utilisée dès la première version d'un prototype fonctionnel. Cette méthode se présente comme une procédure de test utilisateur personnalisable. Dans cette thèse, nous nous intéressons plus particulièrement à deux éléments de cette procédure : un protocole d'élaboration de tâches de RE et une grille d'analyse d'enregistrements vidéo de session de RE. La pertinence du modèle ainsi que les méthodes qui en découlent ont été évaluées à l'occasion de tests utilisateurs. Le modèle, les heuristiques et le protocole d'élaboration des tâches de RE ont été validés. Les premières évaluations de la grille d'analyse d'enregistrements vidéos ont révélé des points à améliorer.

Mots-clefs: Méthodes d'évaluation, Recherche Exploratoire, Méthodes basées sur un Model, Approche centrée utilisateurs

Abstract

Exploratory search systems are search engines that help users to explore a topic of interest. A shortcoming of current evaluation methods is that they cannot be used to determine if an exploratory search system can effectively help the user in performing exploratory search tasks. Indeed, the assessment cannot be the same between classic search systems (such as Google, Bing, Yahoo!...) and exploratory search systems. The complexity and the difficulty to have a consensus definition of the exploratory search concept and process are reflected in the difficulties to evaluate such systems. Indeed, they combine several specifics features and behaviors forming an alchemy difficult to evaluate. The main objective of this thesis is to propose for the designers of these systems (*i.e.* computer scientists) user-centered evaluation methods of exploratory search systems. These methods are based on a model of exploratory search process in order to help the evaluators to verify if a given system supports effectively the exploratory search process. Thus, after elaborating a model of exploratory search process, we propose two model-based methods, with and without users, which can be used all along the design process. The first method, without users, can be used from the first sketch of the system, consists of a set of heuristics of exploratory search and a procedure for using them. We also propose two tools facilitating their use: an online form format and an Google Chrome plugin, CheXplore. The second method involves real end-users of exploratory search systems who test a functional prototype or version of an exploratory search system. In this thesis, we mainly focus on two model-based elements of a customizable user testing procedure: a protocol for the elaboration of exploratory search tasks and a video analysis grid for the evaluation of recorded exploratory search sessions.

Key words: Evaluation methods, Exploratory Search, Model-based methods, User-centered approach

Remerciements

Je remercie mon directeur de thèse Fabien Gandon et mes deux co-encadrants, Alain Giboin et Raphaël Troncy, pour m'avoir accordé leur confiance et pour l'aide apportée tout au long de cette riche aventure.
Merci également à Johan Montagnat et au Labex UCN@Sophia pour avoir accepté de financer ce projet de thèse singularisé par sa nature pluridisciplinaire.

Merci à Nicolas Marie pour m'avoir fait découvrir lors d'un stage de fin de Master passionnant le domaine de recherche qui m'a amenée en thèse.
Merci à Jean-Marie Dormoy, mon ancien stagiaire, pour avoir développé une extension Chrome permettant une utilisation plus aboutie d'une des méthodes proposées.
Merci également à Patrice Pena pour son aide précieuse dans le processus de validation du modèle de recherche exploratoire.

J'adresse mes remerciements à Christian Bastien, Professeur à l'Université de Lorraine, et Pierre De Loor, Professeur à l'Ecole Nationale d'Ingénieurs de Brest, pour avoir accepté d'être les rapporteurs de cette thèse. Je remercie aussi Sylvie Catellin, Maître de conférences à l'Université de Versailles-Saint-Quentin-en-Yvelines, pour avoir accepté d'être l'examinatrice de ma thèse.

Plus personnellement, merci aux amis rencontrés ces dernières années au sein de l'équipe Wimmics, ces formidables compagnons de route qui m'ont apporté un soutien constant et infaillible. Ils ont rendu cette aventure plus humaine, et plus riche encore.
Merci à Amel Ben Othmane, Abdoul Macina et Amosse Edouard pour tous ces bons moments passés ensemble, pour leur écoute et leur bonne humeur.
Merci également à Papa Fary Diallo qui, lui aussi, a toujours été là pour chacun d'entre nous. Où qu'il soit à présent, je me souviendrai toujours de lui, de son amitié chaleureuse et bienveillante. Merci pour tout ce que son passage sur cette terre nous aura apporté.
Merci à Christine Foggia, la fabuleuse assistante d'équipe, pour son aide répétée et pour sa joie de vivre.

Merci à Alexandre Monnin pour sa confiance, ses encouragements et ses nombreux conseils. Nos discussions ont toujours été pour moi une véritable source d'inspiration.

Merci aussi aux autres membres, actuels ou passés, de l'équipe Wimmics, SPARKS ou Data Science d'EURECOM pour leur encouragements et pour les discussions aussi animées que chaleureuses : Molka Dhouib, Mehdi Ahizoune, Sami Lazreg, Tom Bosc, Geraud Fokou, Chihab Bouchenaki, Yoann Bertrand, Michel Buffa, Julien Plu. . .

Merci aux membres de la MSHS Sud-Est pour leur aide, leur bienveillance, pour les déjeuners toujours empreints de bonne humeur et pour les discussions et séminaires passionnants : Pierre Thérouanne, Bernard Conein, Manuel Boutet, Marc Relieu, Lise Arena, et Catherine Félix.

Il est des amitiés qui ne souffrent pas du temps qui passe. Je tiens à remercier ces personnes que la vie a placé sur mon chemin par un heureux hasard, il y a plus de dix ans pour certaines, et qui, depuis lors, n'ont cessé de m'encourager et de me faire sourire. Un grand merci à Anaëlle Hanni, Lesly Lonoh, Noémie Samuel, Emma Toesca, Guillaume Plumat, Morgane Pepe, Louise Chaussade et Camille Amerigo.

Je remercie avec une infinie tendresse mes parents, Emmy et Philippe, qui, depuis toujours, m'ont encouragée à être curieuse et à dépasser mes limites. Merci pour leur confiance, leur soutien constant et leur amour inépuisable.
Merci à mon frère, Alexandre, le premier docteur de la famille, avec qui j'ai pu partager les hauts et les bas de la thèse. Merci pour son amour et sa présence, même à l'autre bout du monde - là où il fait si froid !

> *Success is not final, failure is not fatal: it is the courage to continue that counts.*

— Winston Churchill

A mes parents, Emmy et Philippe

Contents

List of Figures

List of Tables

1 Introduction

> *Savoir s'étonner à propos est le premier pas fait sur la route de la découverte.*
>
> **— Louis Pasteur**

In a context of search engines increasing the personalization of their results (*e.g.* Google, Yahoo! or Bing), exploring and learning new things becomes for the users, as well, increasingly difficult. Eli Pariser introduced the notion of *filter bubble* in [58], which consists of an intellectual isolation that can occur when search engines' or websites' algorithms filter the information they provide, without the knowledge of the users. These algorithms, in the same approach as personalized advertising, base this selection of information on the users' browsing history, or their "*likes*" on social networks (*e.g.* Facebook, Twitter...) for instance. Thus, the provided results are those that suit to users' opinion and preferences, because these personalized information are presupposed more *relevant* for users. This selection process therefore locks the users in a personalized ecosystem of information, a *filter bubble*, created by these algorithms. The main limitation of this filter bubble is that users are progressively shown an increasingly more biased vision of their surrounding world. Indeed, everything is made to please and comfort the users: they are therefore less exposed to contradicting points of view. The users become *intellectually isolated*, and it becomes difficult for them to learn and discover new things.

The concept of filter bubble was especially popularized after the election of the United States president in 2016. Many journalists and people, with liberal political opinion, did not expect the election of Donald Trump as the president of the United States because they were locked in their opinions by a filter bubble which did not provide them with enough information on the Republican supporters[1]. In other words, users having a liberal political orientation had been presented more information with similar political opinion, than information favorable to the opposing party. They did not have a full picture of the situation, but a rather biased one.

For a given subject, impartial discoveries or information monitoring are almost impossible and become more and more crucial in today's world. In the field of

[1] For example, `https://goo.gl/a69kuP` ; `https://goo.gl/mpESJ7`

Computer Sciences, specific search systems are designed in order to support and help users to explore a topic and, or, data without any, or as little personalization of the results as possible. They are called *exploratory search systems*. Such systems aim to support the users' **exploratory searches**.

The main objective of this thesis was to elaborate a set of user-centered methods allowing exploratory search engines' designers (computer scientists) to evaluate their systems. User-centeredness meaning to seek the greater compatibility between the functionalities of a system and the processes carried out by users to perform the tasks to be supported by the system, the methods ought to allow designers of a given system to verify if it effectively supports the exploratory search process.

1.1 Evaluation of exploratory search: an open-ended problem

For many years, almost from the outset, the evaluation of exploratory search systems, *i.e.* how can we say whether these systems support exploratory search process, is one of the principal research focus of the exploratory search systems designers. Indeed, this interest in evaluation of such exploratory search engines leads to:

- the organization of a first workshop in 2006 on the evaluation of exploratory search systems in the context of a conference on research and development in information retrieval [78];
- the organization of a second workshop in the context of a conference on Human-Computer Interaction in 2007 [77];
- as a result of these two workshops, a special issue was published in the *Information Processing and Management* journal in 2008 [76]; and,
- the important chapter "*Evaluation of Exploratory Search Systems*" in the reference book "Exploratory Search: Beyond the Query-Response Paradigm" in 2009 [75] .

Since then, the importance of the evaluation of exploratory search systems did not decreased with the ACM Workshop on Exploratory Search and Interactive Data Analytics (since 2017) where the issue of evaluating such systems is frequently discussed [20, 21]. The researchers agree on the need to address the issue.

1.2 Exploratory search systems evaluation: the need for specific methods

Exploratory search is a specific process or activity which demands to be supported by dedicated systems. By extension, these exploratory search systems need dedicated evaluation methods which aim to verify if they effectively support exploratory search. The existing evaluation methods of exploratory search systems are still incomplete and need:

- A more stabilized definition of the exploratory search characteristics (*e.g.* a description of the users' needs and behaviors in an exploratory search task);
- A more suitable process model of the exploratory search process. Generally the evaluation methods do not use a model of exploratory search process, and if one model is used, the selected model is not a suitable one;
- More suitable metrics. The commonly used metrics are those used in the evaluation of *classic search engines*. They are more focused on the outcomes and the algorithms and it is not sufficient in the evaluation of exploratory search systems. In [75], the authors say that they "may be *inappropriate*";
- A more suitable user-centered procedure. The existing methods do not sufficiently take into account the user interaction behavior.

1.2.1 Exploratory search: an undefined concept

Exploratory search is a particular information seeking activity. It is a loosely defined concept as its definition is not stable and continues to evolve. Marchionini proposes in [33] the first attempt to characterize exploratory search and describes "a set of search activities associated with an exploratory search process" [74] such as knowledge acquisition, comparison, analysis or evaluation. He compares exploratory search to the most basic and well-known type of search activity, *lookup*, which refers to focused searches where the user has a specific goal in mind and an idea of the expected result. The main goal in exploratory search is *learning*. But, "learning in exploratory search is not only about memorization of salient facts, but rather the development of higher-level intellectual capabilities" [74].

The exploratory search concept have no stabilized or consensus definition adopted by the researchers of the field; rather they agree that it is a complex and unclear concept which is still in its infancy.

1.2.2 The existing evaluation methods issues

Designing exploratory search systems that support the exploratory search process involves "highly interactive interfaces" [2]. Mark Nolan says that "*in the design of search results and interfaces for browsing rich information resources we need to design a certain degree of elasticity into the product to give users more control over the results*" [50]. White and Roth add that the evaluation of exploratory search systems needs subjective measures (such as user satisfaction, engagement and information novelty) but it is through the evaluation of interaction behaviors and cognitive aspects that "one can truly evaluate the effectiveness" of exploratory search systems [75].

Then, evaluating such search engines is still an open issue. One of the reasons is that the concept of exploratory search is loosely defined and does not have yet a clear and stable definition. Moreover, the process of exploratory search itself is continuously changing and different depending on the user (expertise, attention, motivation, fatigue...), the context (the location, the noise around...) and the system (the database, the functionalities...). As a result, the evaluation methods of exploratory search systems are still incomplete as they are not fully based on a suitable exploratory search process model.

The issue of evaluating exploratory search systems arises because these specific search engines cannot be evaluated in the same way, metrics or methods as classic search engines that mainly support *lookup searches*. Indeed, the process, the goals of search, or the users' expectations are really different in terms of users' needs, behaviors, or even in terms of interactions between the systems.

A shortcoming of existing evaluation methods is that they cannot be used to determine if an exploratory search system can effectively help the user in performing exploratory search activity. The complexity and the difficulty to have a consensus definition of the exploratory search concept and process are reflected in the difficulties to evaluate such systems. Their evaluation is recognized as a difficult and subtle activity because "it entails a qualitative and quantitative analysis both of the user behavior and of the search results" [12].

1.3 Thesis research questions

In this thesis, we answer to the following research questions:

1. Which model of exploratory search process to choose or to elaborate for designing more appropirate evaluation methods?

2. How to elaborate a model of exploratory search process which will be used as a basis for our methods?
3. How to design model-based methods?
4. How to design designer-oriented methods?

1.4 Thesis contributions: model-based evaluation methods for exploratory search systems

In this thesis, we answer the research questions raised earlier. Our work and the research questions are at the intersection of different research domains: cognitive ergonomics, human computer interaction, cognitive psychology, information seeking and retrieval, computer sciences... Our approach can be seen as top-down: we base our methods on a model of the users' exploratory search process. The main contributions of this thesis are as follows:

- A set of stabilized exploratory search characteristics. This set is used as an analysis grid for the evaluation of existing information seeking models as candidate models on which the evaluation methods to be designed could be based[2].
- A model of exploratory search based on the previous evaluation of information seeking models[3].
- A model-based evaluation method without users: a heuristics inspection method which includes[4]:
  - The heuristics of exploratory search;
  - A procedure of use of the heurisitc: an online evaluation checklist for an easy use of the heuristics of exploratory search;
  - A Google Chrome plugin to facilitate the inspection of exploratory search systems with our heuristics of exploratory search.
- Elements of a model-based evaluation method with users: a customizable user testing procedure including:
  - a protocol for the elaboration of exploratory search tasks;
  - a video analysis grid.

[2]This work was published at the Workshop Exploratory Search and Interactive Data Analytics (ESIDA) 2017 in the context on Intelligent User Interface (IUI) conference [56]
[3]This work was published in the Work in Progress track at IHM conference 2018 [57]
[4]This work was published in the Work in progress track in the context of British Human Computer Interaction (BHCI) conference 2018 [55]

1.5 Thesis Structure

The remainder of this dissertation is as follows:

Chapter 2

In Chapter 2 we provide a survey of the existing exploratory search systems and the ways they are evaluated. We underline the necessity of a user-centered evaluation of these systems. We also introduce our approach to the design of the model-based evaluation methods.

Chapter 3

In Chapter 3 we present our approach to the design of a model of the exploratory search process. We also present our model and how we evaluated its relevance by confronting it to the observed behaviors of twelve users performing a real exploratory search session.

Chapter 4

In Chapter 4, we present our first evaluation method: an inspection method. The method consists of a set of heuristics and a procedure for using them. After presenting the design method of the heuristics of exploratory search, we describe the evaluation of the proposed method.

Chapter 5

In Chapter 5, we present our second evaluation method which consists of a customizable user test procedure. In this chapter we mainly focus on two elements of this procedure: the protocol for the elaboration of exploratory search tasks, and a video analysis grid for the evaluation of recorded exploratory search sessions. For each of these elements, we describe how we designed them and how we assessed them.

Chapter 6

In Chapter 6, we present CheXplore, our Google Chrome plugin. This plugin aims to facilitate the interface inspection of exploratory search systems with our heuristics of exploratory search presented in Chapter 4.

Chapter 7

In Chapter 7, we conclude the thesis by summarizing its main contributions and proposing some perspectives.

2 General state-of-the-art and Approach

In this Chapter we provide a general sate-of-the-art related to the exploratory search systems evaluation issue; further details will be presented in the following chapters. This general context allows us to introduce our approach. Here, we present an overview of the concept of exploratory search, and in a second time we link the difficulty to have a stable definition of this concept to the limits of existing evaluation methods of exploratory search engines. We present our approach to the design of model-based evaluation methods.

2.1 Exploratory search: an undefined concept

Exploratory search is a particular information seeking activity. This loosely defined concept has an unstable definition which continues to evolve. The first attempt to characterize it was [33] written by Marchionini in 2006. Exploratory search refers to cognitive-consuming search tasks like learning or investigation [36], see Figure 2.1. Marchionini compares exploratory search to the most basic and well-known type of search activity, *lookup*, which refers to focused searches where the user has a specific goal in mind and an idea of the expected result. A typical example would be a user wanting to make a reservation to a restaurant and looking for the phone number on the Web. On the other hand, exploratory search is described as open-ended, with an unclear information need (as in Belkin's anomalous state of knowledge [8]), an ill-structured problem of search with multiple targets. This search activity is evolving and can occur over time. For example, a user wants to know more about *France*, she does not really know what kind of information she wants or what she will discover in this search session; she only knows she wants to learn more about that topic: it can be France's history, kings, cheeses, wine... Hence, the main goal in exploratory search is *learning*. But, "learning in exploratory search is not only about memorization of salient facts, but rather the development of higher-level intellectual capabilities" [74]. In addition, the process is continuously changing and depends on various factors such as the user (her expertise on the explored subject, her attention, her motivation, her fatigue...), the context (the physical environment, *i.e.* where the user is performing her search...), and the system (the data base, its functionalities...). In fact, two exploratory searches with the same goal of search

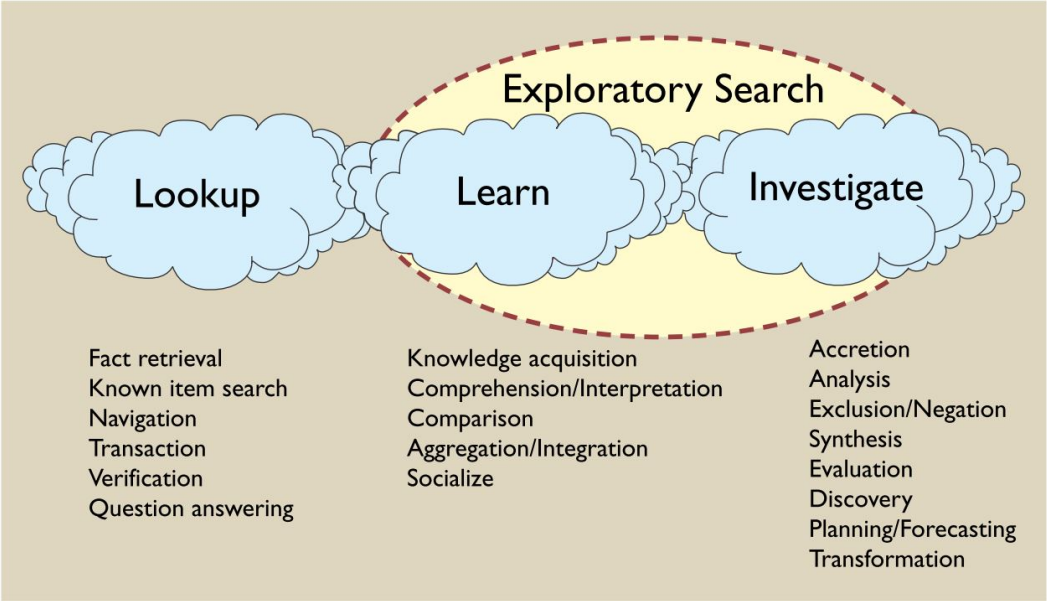


Fig. 2.1: Taxonomy of search tasks proposed by Marchionini in [33]

cannot be identical, and the relevance of a result may vary depending on the user, the time... It is an unpredictable search activity.

Many works use the dichotomy lookup/exploratory searches to define or introduce the exploratory search concept. The exploratory search's complexity is reflected in this approach of contrasting the two types of searches in order to propose a definition of exploratory search concept, and to make it understandable. However, we believe in a continuum/spectruum between these two extremes [56], because we can find some lookup activities in an exploratory search session [75]. Indeed, Marchionini depicts in Figure 2.1 these two search activities as overlapping clouds, suggesting that "lookup tasks are embedded in exploratory tasks and vice versa" [33].

White proposes in [74] a complete definition of exploratory search characteristics: "*[the term] exploratory search can be used to describe both an information-seeking problem context that is open-ended, persistent, and multifaceted, and an information seeking process that is opportunistic, iterative, and multi-tactical. [...] Although almost all searches are in some way exploratory, it is not only the act of exploration that makes a search exploratory; the search must also include complex cognitive activities associated with knowledge acquisition*".

In addition to exploratory search characteristics, this concept is also related in the literature to other information seeking behaviors such as *information foraging*. The information foraging theory [62, 61] attempts to understand and explain how people seek information. The authors relied food foraging behaviors to information seeking behaviors, in a sense that behavioral patterns are similar. For example, based on the

information scent concept, information seekers detect and use cues (e.g. Web links or bibliographic citations [60]) to move from one information path to another, looking for relevant information to their goal. This kind of behavior can be linked to the *informavore* concept which characterizes an organism that consumes information.

Children have exploratory search behaviors when they discover the world around them and learn (new) information about it. They explore without any idea in mind of what they are going to discover everyday. Exploratory search implies a certain level of uncertainty [75] and maybe a little bit of innocence in the way the problems are addressed.

In conclusion, there is no stabilized or consensus definition of exploratory search adopted by the researchers of the field. Actually, they agree in the literature or in conferences that exploratory search is a complex topic of interest, which is still in its infancy. We can retrieved in the literature several characteristics but the process of exploratory search is still poorly researched and little-understood.

2.2 Exploratory search systems and their evaluation: limited methods

Exploratory search is still at an early stage of research. In order to support that specific kind of *information seeking* activity, some systems, called *exploratory search systems*, were designed, mostly in the computer sciences field. Their features and the proposed interactions are essentially many attempts to support exploratory search tasks and process. However, as the exploratory search activity have an unclear definition, the evaluation of exploratory search systems is still an open issue. A main point is the capacity of the evaluation methods to effectively assess whether users' exploratory search behaviors and tasks are actually supported by the exploratory search systems.
In this section we introduce some existing exploratory search systems and the way designers evaluated them.

2.2.1 Exploratory search systems

In the literature, we can identify different kinds of exploratory search systems, with different databases, different interfaces, different user experiences, etc. Table 2.1 presents a non-exhaustive list of existing exploratory search systems.
The main goal of these kinds of system is to support the exploratory search activity. It means that the user's needs when she achieves an exploratory search task are

different from her lookup tasks' needs, and the designers have to elaborate new suitable interfaces supporting this specific activity.

Note that, in Table 2.1 only *Cerchiamo* is a collaborative exploratory search system. The numerical superiority leads us to focus only on individual exploratory search systems in this work.

2.2.2 The need for new evaluation methods of exploratory search systems

In Chapter 1 we said that the evaluation of exploratory search systems is still an open issue and a principal subject of interest of the field of research. This problem is reflected in the designers' difficulty to evaluate their exploratory search systems. In the following subsections we present the importance of exploratory search evaluation and make a survey of the existing methods and their limitations.

2.2.2.1 The importance of exploratory search systems evaluation

Exploratory search is a particular activity and, as mentioned in Section 1.1, its definition is still vague and non-consensual. Evaluating exploratory search systems is, almost from the beginning, one of the principal issue which is still open-ended.

Exploratory search systems aim to support a particular information seeking activity where the goal of search is not really defined in users' mind. They aim to help the users to find and discover what they are not looking for. Actually, this implies some specific features adapted to users' behaviors when they perform an exploratory search task.

Tab. 2.1: Non exhaustive list of existing exploratory search systems (some of them are still accessible online)

Exploratory search systems	Interface	Evaluation	User test	URL	References
SWOC (Semantic wonder Cloud)	Graph	NA	NA	NA	[41]
LED (Lookup Explore Discover)	Tag cloud of query ressources suggestions + list of results	NA	NA	NA	[40]
Aemoo	Form of graph	NA	NA	http://wit.istc.cnr.it/aemoo/	[43]
Seevl	List of artists recommendations	NA	NA	https://developer.seevl.fm/	NA
Yovisto	Result list and query entity suggestion	Quantitative evaluation of heuristics (precision and recall metrics) + Usefulness of the exploratory search feature (user satisfaction)	No + Yes	http://www.yovisto.com/	[72]
Discovery Hub	Carousels (with a top list)	Results' relevance, novelty and unexpectedness (precision, recall,interest and surprise)	Yes	http://discoveryhub.co/	[37, 38, 39, 36]
inWalk	Graph of clusters	NA	NA	http://islab.di.unimi.it/inwalk/	[13]

Continued on next page

Tab. 2.1 – *Continued from previous page*

Exploratory search systems	Interface	Evaluation	User test	URL	References
Linked Jazz	Graph	NA	NA	http://linkedjazz.com/	[59]
3cixty	Parallel interface	System's usability	Yes	https://www.3cixty.com/	[70]
Cerchiamo	3 user interfaces: (1) *Media Magic* (queries and browsing results), (2) *A rapid serial visualization result browsing interface*, and (3) *a shared display*	NA	NA	NA	[22]
Frankenplace	World map	NA	NA	http://frankenplace.com/	[1]
(no named exploratory search system)	Result list (top list)	Relevance of the results (precision, recall and F-measure)	No	NA	[44]
EXPOSE	3 columns layout: a search box panel, an event directory panel and a preference tab	NA	NA	NA	[42]
EVELIN	Result list and a graph visualization	NA	NA	NA	[69]

Tab. 2.1: Non exhaustive list of existing exploratory search systems (some of them are still accessible online)

In their book [75], White and Roth say that "*when evaluating exploratory search systems, it is impossible to completely separate human behavior from system effects because the tools are so closely related to human acts, they become symbiotic*". It means that exploratory search activity must to be supported by features, and more broadly systems, which allow and support this kind of activity in which "*the information-seeking process itself is just as important – if no more so – than the final state*" [26].

The main goal in evaluating exploratory search systems is to determine if they effectively support the users' exploratory search activity and tasks. The evaluation of exploratory search systems is recognized as a difficult and subtle activity because "*it entails a qualitative and quantitative analysis both of the user behavior and of the search results*" [12]. These complex systems combine several functionalities and behaviors forming an "alchemy difficult to evaluate" [54]. The assessment cannot be the same between classic search systems and exploratory search systems.

The question that always arises is the pertinence of the methods and metrics used in the evaluation of exploratory search systems.

2.2.2.2 Survey of existing evaluation methods and their limits

While many exploratory search systems have been described in the literature, the way their designers assess them is very limited and only few of them report evaluations involving real end-users (see Table 2.1). Exploratory search systems are often not evaluated, e.g. [1, 13, 22, 41, 40, 42, 43, 53, 59, 69]. When an evaluation is performed, it generally focuses on one aspect only. It can be, for example, metrics that assess the outcomes such as the standard Information Retrieval *precision* and *recall* metrics [36, 44, 72], or the evaluation of the interface's usability [70]. However, standard Information Retrieval metrics are mainly focused on the result ranking or the algorithm efficiency and they suppose a precisely identified search target and result set. They are not sufficient because the success of exploratory search systems does not only depend on the search algorithm quality [63, 75]. In addition, usability issues may only reveal a bad design by a misunderstanding of the exploratory search system' interface. But such an evaluation excludes the interactions and users' specific needs in an exploratory search task. Then, these evaluations are not sufficient because the success of an exploratory search system mainly depends on an appropriate interaction between the user and the system, so a further characterization of the interaction is necessary. Indeed, in [75], the relation between the user and the system is considered as "intentionally symbiotic", which means that for a complete exploratory search systems the user, as like the system's algorithms, must be taken to account.

Mark Nolan said in [50] that "*in the design of search results and interfaces for browsing rich information resources we need to design a certain degree of elasticity into the product to give users more control over the results*", which involves highly interactive interfaces [2]. Limiting the assessment to one aspect only, and ignoring the importance of the user's exploratory behaviors, do not allow to demonstrate that the systems effectively support the exploratory search process.

Evaluation methods of exploratory search systems are always based on an exploratory search process model, whether implicit or explicit. We only consider the evaluation methods based on explicit models. We observe that the models used in these methods do not exactly reflect the exploratory search task. We provide two examples that illustrate this state of affairs.

In [80], Wilson et al. combine two models in order to automate the evaluation of advanced search interfaces: (1) Bates' model [7] of the four levels of search strategies (move, tactic, stratagem, and strategy), and (2) Belkin's model of information-seeking strategies [9]. Wilson et al. also designed an inspection evaluation method that can be performed without users. However, this evaluation does not take into account the user's exploration process in its entirety. Indeed, it mainly focuses on the tactics the users may employ, and Bates' tactics are low-level actions such as select, focus, check, or pinpoint.

In [12], Bozzon et al. extend Kuhlthau's Information Seeking Process model [28] and use it in their evaluation. Here again, the evaluation does not take into account the user's exploration process in its entirety. Even if the authors involve users in their evaluation, the main problem is the selected model. The model identifies five cognitive or mental phases: *Initialization, Selection, Exploration, Formulation, and Collection*. Exploration is only a component of the search process while exploratory search must be considered as the main activity of the user and not a fraction of the process [75]. Thus, the model cannot reflect the exploratory search process. Another issue is related to the finite-state automaton (FSA) representation of the Kuhlthau model used by the authors. In this representation, the cognitive phases of the model correspond to states, and user actions within the exploratory UI correspond to transitions between states. Here again, the actions are analyzed at a low level: *click, select, zoom*... All the possible actions are listed, for example: *Click on the Select source button for selecting the service to be used as source for the first query*. This low-level actions analysis does not link the actions to higher-level activities. This gap cannot actually allow a complete assessment of an exploratory search system in order to verify if it supports users' exploration behaviors.

To sum up, these two previous model-based evaluation methods propose to analyze the exploratory search process at a too low level and they do not link the considered

basic actions to higher-level activities. This gap prevents from having a complete exploratory search process reflection. In a word, none of these approaches takes into account the user's exploration process in its entirety. To design our model-based evaluation methods of exploratory search systems, we need a model that describes the information seeking process at a *behavioral level*.

2.2.3 Conclusion

The complexity of the task and the difficulty to have a consensus and stable definition of the exploratory search concept are reflected in the limits of existing evaluation methods of exploratory search systems. These evaluation methods can hardly be used to determine if exploratory search systems can support users' behaviors and exploratory search tasks. Indeed, most of the chosen traditional information retrieval metrics ignore the nature of the exploratory search process. Furthermore, the few existing evaluations based explicitly on a model proceed at a too low level avoiding the analysis of a user's exploration process in its entirety. Thus, the evaluation methods of exploratory search are still incomplete as they are not fully based on a suitable exploratory search process model. Indeed, these methods rely on a model of exploratory search which is still loosely defined, or at least on a definition which is not yet clear and stable.

We need a better understanding of the exploratory search process in order to be in position to verify if exploratory search systems effectively assess users' exploratory search behaviors and tasks. Moreover, we need evaluation methods that can be applied easily or, ideally, tools that help and support evaluators in their system assessment.

2.3 Our approach to the design of exploratory search systems evaluation method

The main objective of this thesis is to elaborate a set of user-centered evaluation methods. Our approach can be seen as top-down in the sense that the methods we propose are based on a model of users' exploratory search process. Figure 2.2 depicts our approach schematically.

The first step of our approach is to select an existing model of information seeking which suits the exploratory search process, or to adapt it, or even to elaborate a new model of exploratory search process, the most important being that the model reflects the actual behaviors of users performing an exploratory search task. The

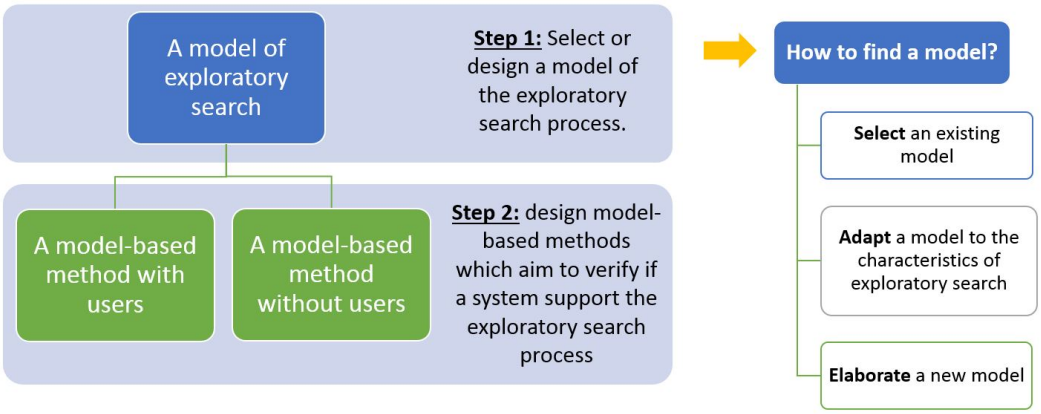


Fig. 2.2: Our top-down approach in the elaboration of exploratory search systems evaluation methods

methods we intend to design are based on this model of exploratory search process, because exploratory search systems have to be suitable to this specific information seeking activity. Then, these model-based methods aim to determine if the evaluated system effectively supports the exploratory search process. The targeted methods are divided into two: one method with users and another without users. The idea is to allow the evaluators to have a set of methods which can be used iteratively all along the design process of an exploratory search system. The method without users should allow preliminary assessment and be less burdensome. It translates the model into concrete interface elements or possible actions required to effectively support exploratory search behaviors. The method with users involves an analysis of users' activity and behaviors in an exploratory search session with a simplified version of the model. Furthermore, the two methods are complementary because they do not highlight the same issues. These targeted methods of exploratory search systems are designer-oriented and discount. In more concrete terms, the methods can be widely and easily used, which means that evaluators will not have to know and understand the model of exploratory search process to use the evaluation methods. In that sense of facilitating the evaluation process, we also want to increase the test automation.

In brief, the main specification of these evaluation methods of exploratory search systems are:

- Two model-based methods which aim to determine if the evaluated systems effectively support the users' exploratory search process:one method with users and another without users.
- Designer-oriented methods, which means:
  - easy procedures;
  - discount methods;

– Tools facilitating the evaluation process.

In the next chapter we present a model of exploratory search which will be used as a basis of two exploratory search systems evaluation methods.

3 A model of exploratory search process

3.1 Introduction: the need for an approximate model

The difficulty to assess exploratory search systems is related to a lack of knowledge on the exploratory search process (Chapter 2). A better understanding of this specific search process is necessary to verify if the exploratory search behaviors are supported by these systems.

The information seeking process model of exploratory search we are looking for is what Donald Norman calls an *approximate model* [51]: it does not depict the whole exploratory search process in detail or precisely, but it is *"good enough for the purpose to which [it will be] applied"* [52]. Norman's model of the seven stages of action is such a model. In [51], Norman claims that the process of performing and evaluating an action can be approximated by seven stages of user activity. He adds that "real activity does not progress as a simple sequence of stages" and the identified stages appear out of order and, in the process, "some may be skipped, some repeated". Unlike the models used in the methods described in Section 2.2.2.2, Norman describes an approximate model process of action at a general level:

- Establishing the Goal
- Forming the Intention
- Specifying the Action Sequence
- Executing the Action
- Perceiving the System State
- Interpreting the State
- Evaluating the System State with respect to the Goals and Intentions

In our case, the model we are interested in should be designed so as to be accurate enough to support the elaboration of evaluation methods of exploratory search systems, and not to describe the whole process of exploratory search exactly. The

model may describe the main actions, at a behavioral level, of the exploratory search process.

Furthermore, in the selection of information seeking models, we choose to restrict our interest to models that describe user's behavior and the information seeking process in a search task.

There are three main ways to get the exploratory search process model we are looking for: (1) we choose an information seeking model that already exists and conform to the description of exploratory search; (2) we choose a model close to the concept and adapt it; (3) we elaborate a new model.

The first difficulty in this search for an exploratory search model is the lack of knowledge we had on the exploratory search process. We cannot find a model if the definition of this search task is still unstable. Thus, we first needed to stabilize the definition of the exploratory search by emphasizing its characteristics. These characteristics allow us to know what to look for in existing information seeking models. We develop this aspect in Section 3.2. In Section 3.3, we present the exploratory search model we designed. In Section 3.4 we provide a preliminary evaluation of the model which leads us to extend it. In Section 3.5, we present the second assessment of the extended model.

3.2 The information seeking models analysis

Even if the definition of exploratory search is still unstable and open ended, some characteristics are frequently found in the literature and the different descriptions of this specific search activity. A certain stabilization of the concept of exploratory search is necessary to know what to look for in the existing information seeking models. We will use these exploratory search characteristics as a grid for evaluating the information seeking process models. To be selected as the model which will underly our evaluation methods, the candidate models will need to be compliant with the acknowledged characteristics.

3.2.1 Characteristics of exploratory search

We started to identify the characteristics of exploratory search that appear most often in literature. Most of the characteristics come from the same few references: [33, 75, 79]. Even though the same characteristics can be found elsewhere under different descriptions, they refer to the same idea. We summarized and listed these characteristics in Table 3.1 [56]. In this table, the order of characteristics is not

meaningful. All are equal insofar as the whole characteristics reflect the exploratory search behavior. Furthermore, some of them are linked, e.g. characteristics 5 and 6 do not refer to the same thing but they share the same definition because they are linked. Moreover, some of them refer to other characteristics, e.g. the "evolving process" characteristic, refers in its definition to characteristics 3, 4, 5, 6 and 7.

Tab. 3.1: Eleven characteristics of exploratory search

Characteristics	**Definition**	**References (non exhaustive list)**
1. An evolving search process	The user adopts an opportunistic behavior, and will change or specify the objective or goals of search or even the strategies used to achieve them through multiple queries reformulation or refinement. During the search, the user can accomplish forward or backward steps.	[36, 44, 54, 74, 75, 79]
2. Several one-off pinpoint searches	Throughout the search session, the user can do several one-off pinpoint searches, e.g. she's looking for a specified information to better understand a result or the reason why the result was proposed. These pinpoint searches can be related to the exploratory search task or not. This is closely related to sensemaking activities.	[33, 75, 74]
3. An evolving information need	Throughout the search session, the user has an evolving information need. The elements or results discovered may change her information need and the way she first considered the framework of the search. This evolution of the information need may appear several times in one search session. It is closely related to characteristic n°1.	[39, 75, 74, 79]

Continued on next page

Tab. 3.1 – *Continued from previous page*

Characteristics	**Definition**	**References (non exhaustive list)**
4. Multiple targets/ goals of search	The user may not have one single precise goal, but rather one vague objective and several smaller goals which may change or evolve during the exploratory search task so as to achieve it.	[4, 36, 39, 54, 74, 75]
5. Multiple possible answers	As the user has one vague objective and several smaller goals to achieve it (see characteristic n°4.), the user might not have one precise answer but an aggregate of relevant information which will help her go further in her reflection and exploratory search process.	[4, 36, 39, 64, 75]
6. Not an expected exact answer		
7. A serendipitous attitude	Having a serendipitous attitude is the faculty to be surprised and to pay attention to it. The user carries out her search by adopting a serendipitous attitude; with such open mindedness, she can allow herself to be surprised by some unexpected element. She then exploits this discovery by changing the search strategy or search goal/objective, etc.	[33, 64, 74]
8. An open ended search activity which can occur over time	The user might never end her exploratory search. She can stop it for multiple reasons (she considers she has enough information to perform another task for example; she doesn't have time to carry on the search; etc.), and she will continue the search few hours/days/weeks/months/years later.	[33, 39, 54, 75, 74, 79]

Continued on next page

Tab. 3.1 – *Continued from previous page*

Characteristics	**Definition**	**References (non exhaustive list)**
9. An Anomalous State of Knowledge (ASK) and an ill-structured (vague, general or unsure) context of search or goals	At the beginning, the user has an ASK and a general context of search: she knows the motivation to start the search, but does not have a precise idea of what she is actually looking for (type of results, kind of information). She only has a lack of knowledge, a vague objective of search but no specific of definitive plan to attain it.	[36, 54, 75, 74, 79]
10. Multifaceted	During the exploratory search, the user selects one or multiple filters or facets, to explore the information space. She will try to find an approach to her problem, she may find an angle of attack or a framework which may include these facets of the explored subject.	[4, 36, 75, 74, 79]
11. Uncertainty is fluctuating	The user starts the search with an intense feeling of uncertainty. The level of uncertainty is intrinsically linked to the specification of the problem. The further the user goes in her search tasks (she will specify her objective and maybe define an approximate plan), the more she reduces her uncertainty. But if somewhere along the way she changes her objectives, the uncertainty will tend to increase again.	[36, 74, 79]

Tab. 3.1: Eleven characteristics of exploratory search

Contrary to the notion of "serendipity", the notion of "serendipitous attitude" that describes the user's state of mind, cannot be found in the literature. Nevertheless, this idea or concept, matches other descriptions / characteristics such as serendipitous

discoveries or opportunistic behavior. We were inspired by Sylvie Catellin's and Forster and Ford's descriptions of the serendipity concept [19, 14].

3.2.1.1 Exploratory search and serendipity: two related but different concepts

In the literature, exploratory search is often linked and sometimes confused to another concept: *serendipity*, such as in [19, 23, 36, 72, 73, 74, 75]. Moreover, [24, 27] add serendipity to the exploratory search's component, proposed by Marchionini, assuming that "*exploratory search behavior comprises a mixture of serendipity, learning, and investigation*". These elements show that the line between these two concepts is blurred, which can lead to a certain confusion.

Horace Walpole created the word *serendipity* in 1754. He based his neologism on the tale of The Three Princes of Serendip and the camel. Walpole's definition of serendipity is the human's capacity to discover by accident and sagacity [14]. Contrary to some common definition, serendipity is not only the act of finding by chance what one is not looking for. Indeed, the role of human's surprise, interpretation and sagacity, which is the acuteness of mental discernment and soundness of judgment, are a part of serendipity concept which cannot be avoided in its definition. Indeed, the human's cognitive processes and the element encountered by chance cannot be separated in a *serendipitous discovery*. Unfortunately, the second and impoverished description of the concept is the most widespread, especially in English-speaking countries or even in the information seeking or retrieval domain [3, 15, 74].

Serendipity may occur in many domains and is widely regarded as valuable in their process. Indeed, important discoveries in sciences (experimental, humanities and social sciences, biology...), art or even daily life for example are linked to serendipity. The most important aspect of serendipity is that it cannot be induced [14]. Indeed, we can create a favourable environment for its apparition, but serendipity is linked to uncontrollable aspects, *e.g.* we cannot induce human's disposition to make such discoveries.

We focus here on serendipity in the context of information seeking or retrieval. When a user performs a search on a search engine, she can make serendipitous discoveries. André et al. say that these discoveries in a search session potentially match search results that are interesting but not highly linked to users' goal of search [3]. This is really interesting because we can find similar aspects in the definition we made of exploratory search results in previous work [39]. We describe exploratory search results in two axes, their interestingness and surprisingness:

- a result is interesting if:
  - the user thinks it is similar to the explored topic, *i.e.* if the user finds a result relevant to her goal of search;
  - the user thinks she will remind or reuse it, *i.e.* if she finds the result interesting but not highly relevant to the actual search.
- a result is surprising if:
  - the user discovered an unknown resource or relation, *i.e.* between the query and a result or between two results;
  - the user discovered something unexpected.

Serendipitous discoveries are undeniably linked to a certain kind of exploration and exploratory searches. We can find similarities between the one who makes serendipitous discoveries and children's exploratory behaviors mentioned in Section 2.1: both are curious of the surrounding world, intuitive, creative, with an intrinsic motivation to explore and discover new things [14, 71].

Even if serendipity can be an important part of the exploratory search process, exploratory searches do not always lead to serendipitous discoveries (users can find relevant information for their search which are not always serendipitous discoveries). As mentioned earlier, we cannot induce serendipitous discoveries, but we can design exploratory search systems that support the exploratory search process which, consequently, will be more favorable to discoveries, whether serendipitous or not.

3.2.1.2 Categorisation of the exploratory search characteristics

Based on the previous descriptions of the exploratory search characteristics, we observed that they can be clustered into three categories (see Figure 3.1) [56]:

- The *User* category refers to characteristics which are related to the inner state of the user (feelings, attitude, expectations, etc.).
- The *Exploratory search task* category refers to characteristics which are related to the exploratory search task and its process, e.g. the description of the task or strategies employed by the user.
- The *User & Exploratory search task* category which refers to characteristics under both the user and the exploratory search task

The characteristics of these three categories are presented in Figure 3.1. It also depicts the imperative need to take into account the user and her task in the exploratory search system development process: from the design to the assessment.

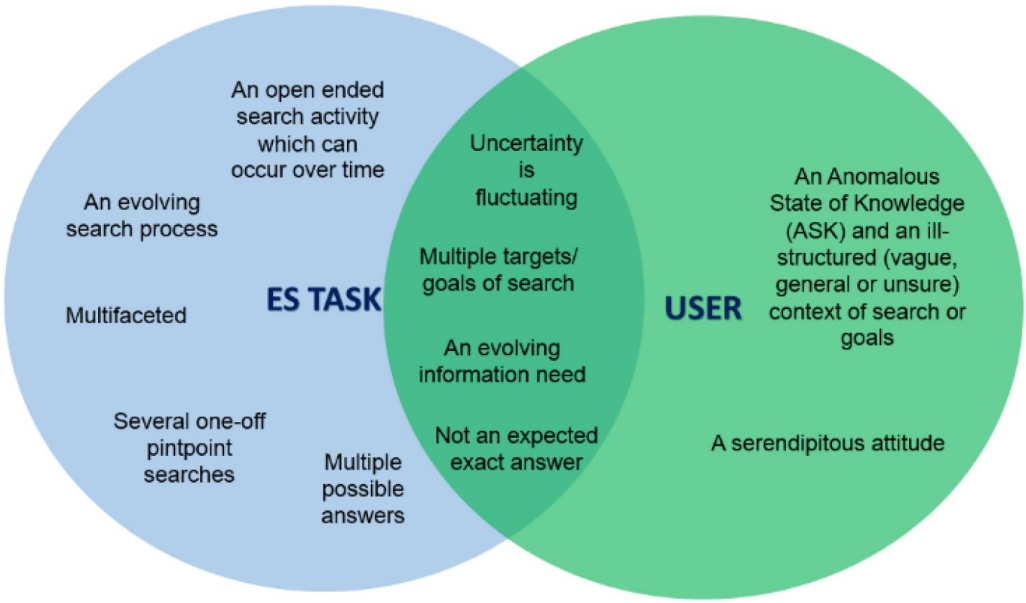


Fig. 3.1: Venn diagram of characteristics of exploratory search.

We have to understand the user's specific needs and the task she performs to adapt and improve exploratory search systems.

3.2.2 Existing models of information seeking

It is difficult to differentiate information seeking models and information retrieval models, especially interactive information retrieval models and information seeking models involving search component and process. Xie distinguishes in [81] two types of models, regardless of what terms are used by their authors (*information retrieval* or *information seeking*): models that mainly illustrate the information seeking process, and models which emphasize the factors influencing the process. In our selection of models, we choose the first type of models because we want a model that takes more into account the user and her behaviors in the search process. These process models represent generally searchers' activities in a "multi-stage representation" [74] and many of them were developed before the appearance of the Web or at it early stage. However, they present relevant aspects of information seeking behaviors that "transcend technological advance" [74]: they describe typical information seeking activities such as "*clarify vague information needs, learn for information present in the collection, and investigate solutions to information problems*" [74].

We select five candidate models that suit to this definition: Ellis model in its extended version [17, 18], Bates' berrypicking approach [6], Kuhlthau's information seeking

process model [28], and Marchionini's model of information seeking [34] and Marchionini's model exploratory search [33].

3.2.2.1 Ellis' model

Based on empirical studies, Ellis' model proposes a set of six types of information activities, or *features*, characterizing the information seeking patterns of real information seekers [17]. Few years later, the model was extended in [18] with two other features. Ellis' features are described as follows:

- *Starting* activities such as the initial search for and overview of the literature or locating key people working in the field;
- *Chaining:* following footnotes and citations in known material or "forward" chaining from known items through citation indexes or proceeding in personal networks;
- *Browsing:* variably directed and structured scanning of primary and secondary sources;
- *Differentiating:* using known differences in information sources as a way of filtering the amount of information obtained;
- *Monitoring:* regularly following developments in a field through particular formal and informal channels and sources;
- *Extracting:* selectively identifying relevant material in an information source;
- *Verifying:* checking the accuracy of information;
- *Ending:* activities actually finishing the information seeking process.

These features form a framework for information seeking [75]. They provide a "*framework for a flexible model to underpin recommendations for information retrieval system design and evaluation*" and they can be employed to derive a set of general recommendations [17]. The model does not specify the order in which the features are carried out: an information seeker is not "guaranteed to undergo an identical information-seeking process as outlined in the model" [75]. The model does not define either the interactions or interrelationships between the features. Ellis explains that "*the detailed interrelation or interaction of the features in any individual information seeking pattern will depend on the unique circumstances of the information seeking activities of the person concerned at that particular point in time*". The model explicitly describes fairly concrete process steps [25]. However, the freedom in this search representation is really relevant for the design of our exploratory search model.

Most situations involving information seeking can be characterized by this model, which might include exploratory search. Moreover, Ellis identifies the model as a flexible behavioral model rather than a process model.

3.2.2.2 Bates' model

Bates' berrypicking model is much closer to the real behavior of information searchers than the traditional model of information retrieval [6]. It highlights the dynamic nature of the information seeking process [11]. It is one of the first model which underline the exploratory nature of this activity. In the model's description, the author proposes a dynamic, nonlinear and evolving search process.

The term *berrypicking* is an analogy to "*picking berries in a forest; berries are scattered on bushes, not in bunches, and the seeker must pick the berries singly*" [75]. The searcher moves into the information space to find relevant information, one by one, dispersed into several documents. All along the search, every step gives new ideas to the searcher and can redefine the query or the search goal (see Figure 3.2). It's a constant renewal of the information need.

Berrypicking model differs from classic lookup model in two aspects. The first is the nature of the query which is described as an evolving one: the queries change and evolve during the search process. The second concerns the search process which follows a berrypicking pattern: each stage consists in the identification of a relevant information which leads to a modification of the search (e.g. a new query). In other words, Bates highlights the importance of the information encountered in the modification of the information need.

Boubée and Tricot claim in [11] that Bates' berrypicking model offers, in that sense, a formalization of the serendipity concept. We are not agree with this claim. Indeed, the berrypicking model illustrates exploration behaviors, like the opportunistic behavior or evolving search. Nevertheless, White and Roth describes berrypicking as a "commonly used strategy in exploratory searches" [75] (which can be common with serendipity); and there is no mention in Bates' model of the *surprise* linked to the information encountered, which is essential to the serendipity concept.

3.2.2.3 Kuhlthau's model

Based on a series of empirical studies, mainly with high school students, Kuhlthau developed a model of information seeking behavior from the user's perspective. Here,

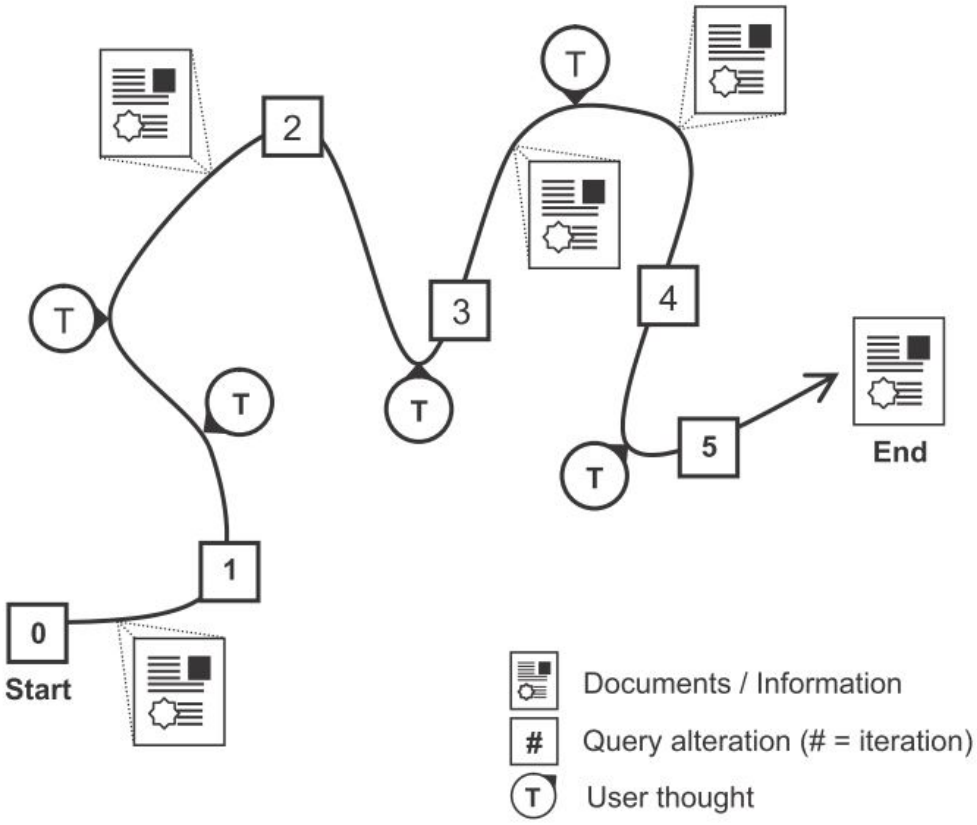


Fig. 3.2: Berrypicking model proposed in [6], (picture from [75]).

information seeking is seen as a problem solving activity [34]. Her Information Seeking Process (ISP) model is explicitly a process model and proposes several stages of a successful search [11]. Kuhlthau's model incorporates the psychological aspects of search [75] and proposes that the feelings of doubt, anxiety, and frustration play a role in the search process [74]. Kuhlthau's approach of the search can be considered, for this reason, as *holistic*, with a more global description of the information seeking process [11].

Kuhlthau describes a model with six stages and incorporates, like nobody did before, three realms common to each stage: the affective (feelings), the cognitive (thoughts), and the physical (actions) [28]. The ISP stages are defined as follows by[25]:

1. Initiation: becoming aware of the need for information, when facing a problem
2. Selection: the general topic for seeking information is identified and selected
3. Exploration: seeking and investigating information on the general topic
4. Focus formulation: fixing and structuring of the problem to be solved
5. Collection: gathering pertinent information for the focused topic
6. Presentation: completing seeking, reporting, and using the result of the task

Stages in ISP	Feelings Common to Each Stage	Thoughts Common to Each Stage	Actions Common to Each Stage	Appropriate Task According to Kuhlthau Model
1. Initiation	Uncertainty	General/ Vague	Seeking Background Information	Recognize
2. Selection	Optimism			Identify
3. Exploration	Confusion/ Frustration/ Doubt		Seeking Relevant Information	Investigate
4. Formulation	Clarity	Narrowed/ Clearer		Formulate
5. Collection	Sense of Direction/ Confidence	Increased Interest	Seeking Relevant or Focused Information	Gather
6. Presentation	Relief/ Satisfaction or Disappointment	Clearer or Focused		Complete

Fig. 3.3: Khulthau's model of Information Seeking Proces (ISP), proposed in [28]

The search activity is a dynamic process led by feelings interacting with thoughts and actions (see Figure 3.3). The changes in feelings, thoughts and actions depend on stages. From the initial stage to the end of the search process, the user makes choices to achieve her goals by a complex interplay between these three realms. The importance given to this interplay is really interesting because in exploratory searches, the user must deal with her feelings (e.g. uncertainty), her thoughts (e.g. fairly vague) while she is interacting with exploratory search systems.

3.2.2.4 Marchionini's model (95)

Marchionini's model focuses on a searching process in electronic environments [34]. In this description of the process, the user is central, she is the one who defines the task, interacts with the search system of her choice and determines when the information seeking process is over. The user "possesses unique mental models, experiences, abilities, and preferences" which have an impact on the search process and its development. The experience refers to the expertise level of the user about the setting, the search domain or the chosen system.

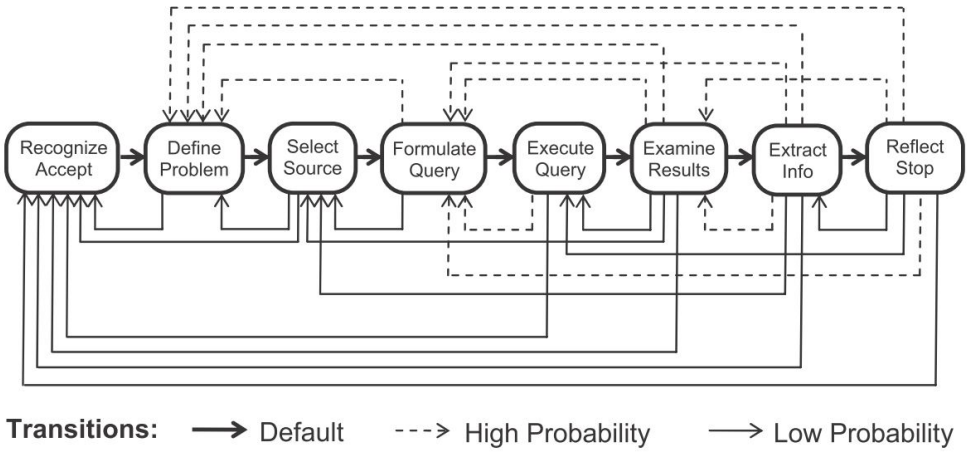


Fig. 3.4: Marchionini's model proposed in [34] (picture from [75])

The model details eight sub-processes which evolve in parallel during the search process:

1. Recognize and accept an information problem
2. Define and understand the problem
3. Choose a search system
4. Formulate a query statement
5. Execute search
6. Examine results
7. Extract information
8. Reflect/iterate/stop

This model is more suitable for electronic environments than Ellis' model [75]. It focuses on "search for information and does not consider learning and understanding" [74], which are two main important activities of exploratory search [33].

The model describes, maybe in great detail the information seeking process by taking into account many elements [11]. Marchionini says that information seeking "*depends on interaction among several factors: information seeker, task, search system, domain, setting and search outcomes*" [34]. The setting is the physical, social and cognitive environment in which the search takes place. The task here, is explicitly goal-driven and the user interact with the system with mental and physical actions. In this model the goal of the search is the same from the beginning to the end of the search, but the task evolves throughout the search process. Figure 3.4 of [34] shows an evolving process with many possible transitions between the sub-processes. There are three types of transitions : *default, high probability* and *low probability*.

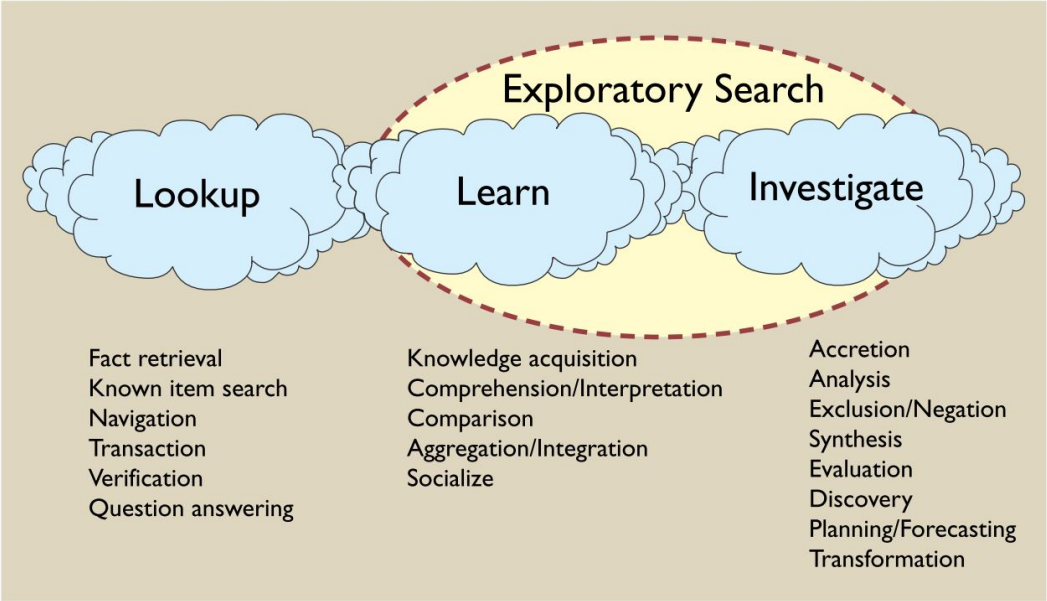


Fig. 3.5: Taxonomy of search tasks proposed by Marchionini in [33]

3.2.2.5 Marchionini's model (06)

The description of exploratory search in [33] is the first attempt for exploratory search tasks characterization. This seminal model is often used by authors working on exploratory search.

Marchionini compares exploratory search to the most basic and well-known type of search activity, *lookup*, which refers to focused searches where the user has a specific goal in mind and an idea of the expected result. Marchionini depicts these two search activities as overlapping clouds (Figure 3.5), suggesting that "lookup tasks are embedded in exploratory tasks and vice versa" [4].

Marchionini's model proposes among other things a set of activities, based on Bloom's taxonomy [10], related to exploratory search and highlights the activities associated to exploration (learn and investigate) or lookup: such as *knowledge acquisition, comparison, analysis or evaluation* for exploratory search, and *known-item search, verification and fact retrieval* for lookup. This model presents the interplay between these activities (and sub-activities) but does not detail them.

The main goal in exploratory search is learning. But, "learning in exploratory search is not only about memorization of salient facts, but rather the development of higher-level intellectual capabilities" [74]. Thus, Marchionini's model describes the exploratory search activity at an intellectual level derived on Bloom's taxonomy of educational objectives [75].

3.2.3 The information seeking model analysis based on the characteristics of exploratory search

We identified in the last sections the characteristics of exploratory search and the information seeking models we want to evaluate.

The next stage is to analyze the identified information seeking process models and confront the author's description of the corresponding model to our list of characteristics of exploratory search. Although we can conceive that any model would not completely fit the exploratory search definition, the idea here is to emphasize models that best satisfy the exploratory search characteristics. If we can adapt one of them to match with the exploratory search characteristics, it may be an imperfect or incomplete exploratory search model, but good enough to help us in the design of an evaluation method of exploratory search systems.

In the analysis, we are continuously looking for characteristics of exploratory search for each information-seeking process model using Table 3.1 as an analytic grid. For a selected model, we check if the characteristic:

1. is explicitly mentioned in the description provided by the author(s); or
2. can be inferred from the description; or
3. is absent or cannot be inferred.

The model we are looking for presents in its description all the exploratory search's characteristics listed previously and depict a model with steps. It also allows a certain freedom in its search process description because exploratory search is an unpredictable one, and a search process can differ among exploratory search systems. Note that the specification of the inferences may show the possibility we will have to adapt the model, in order to have a model which covers all the exploratory search's characteristics. Results of the checking are reported in Table 3.2; in this table, *case 1* is coded as Yes (the characteristic is present), *case 2* is coded as Yes (Inferred), and *case 3* is coded as No (the characteristic is absent).

Neither of the models checked all the exploratory search characteristics of our analytic grid. As we can see, Marchionini's model (06) is the one which fulfills most of the criteria. Nevertheless, this model cannot be referred as a sequential model, or in line with Norman's approximate model. It does not correspond to the multi-step process we are looking for, and we need lower-level activities to mobilize in our method in user tests [56]. It is moreover impossible to adapt it in this way. This is an important point, because we need a model that identifies process *steps* for our evaluation method.

Exploratory Search Characteristics	Ellis	Bates	Kuhlthau	Marchionini 95	Marchionini 06
1. An evolving search process	Yes	Yes	No	Yes	Yes
2. Several one-off pinpoint searches	No	Yes (inferred)	No	No	Yes
3. An evolving information need	No	Yes	No	Yes	Yes
4. Multiple targets/goals of search	No	Yes	No	Yes	Yes (inferred)
5. Multiple possible answers	Yes (inferred)	Yes	No	Yes	Yes (inferred)
6. Not an expected exact answer	Yes (inferred)	Yes	No	No	Yes (inferred)
7. A serendipitous attitude	No	Yes	No	Yes	Yes
8. An open ended search activity which can occur over time	No	No	No	No	Yes
9. An Anomalous State of Knowledge (ASK) and an ill-structured (vague, general or unsure) context of search or goals	Yes (inferred)	Yes (inferred)	Yes	Yes	Yes
10. Multifaceted	No	No	No	Yes	No
11. Uncertainty is fluctuating	No	No	Yes	No	No

Tab. 3.2: Information seeking models analysis.
In this table, (1) "Yes"refers to a characteristic explicitly mentioned in the description provided by the author(s); or (2) "Yes (inferred) refers to a characteristic which can be inferred from the description; or (3) "No" refers to an absent characteristic or a characteristics which cannot be inferred.

Bates' model describes more a strategy used in exploratory searches than exploratory searches activities. Indeed, it is representative of a form of exploratory search behavior but does not detail actual activities a user performs in an exploratory search task (e.g. *formulate a query, analysis of the results or results list...*) [75]. Morever, it does not check all the exploratory search characteristics. As the preceding model, it cannot be used as a basis for a new model.

Kuhlthau's and Marchionini's (95) models describe models with steps, closer to the model we are looking for. As mentioned previously in Section 2.2.2.2, in Kuhlthau's model, exploration is only a component of the search process while exploratory search must be considered as the main activity of the user and not a fraction of the process [75]. In addition, it describes only a successful search process. It means that the user achieves her goals of search. The concept of successfulness is different in exploratory search, because the goals are vague and evolving. A successful exploratory search is more a formation of an aggregate of relevant information to pursue the search or knowledge acquisition. For its part, Marchionini's model (95) describes a complex model, maybe too much detailed, that takes into account too many elements [11]. Even if the description is closer to exploratory search characteristics than Kuhlthau's description, the exploration is again only a part of the search process and not the main activity. Furthermore, as described in Section 3.2.2, the two main activities of exploratory search, learning and understanding, are not fully considered in the process description.

Ellis' model does not fulfill either all the characteristics of exploratory search but its definition and features can be easily adapted to the exploratory search concept and its characteristics. Indeed, it proposes a non-linear process without predefined sequences, and we want a model describing an unpredictable process. The author specified that the model's features can be employed to derive a set of general recommendations. This model was already used as a basis for the design of user-centered evaluation methods, e.g.[32]. Thus, Ellis' description matches our objective of an model-based evaluation method for exploratory search systems. For all these reasons, we selected the Ellis' features-based model of information seeking as a first approximation for the design of our model.

However, noticing that the Ellis' model, in its original form, does not define either the interactions or interrelationships between the features, we needed to adapt the model to better suit the exploratory search concept and its characteristic, leading thus to a second approximation.

3.3 The design method of the model of exploratory search

As mentioned earlier, the selected Ellis' model needs to be adapted to better suit exploratory search concept and its characteristics. Indeed, the features of Ellis' model in its original form does not fulfill all the exploratory search characteristics:

- Starting (char. 9)
- Chaining
- Browsing
- Differentiating(char. 10)
- Monitoring (char. 8)
- Extracting(char. 5)
- Verifying (char. 2 partially)
- Ending (char. 8 partially)

We adapted Ellis' model and designed a new model of exploratory search process.

In the following description of the model, each feature is linked to the corresponding exploratory search characteristic(s) [56] described in Table 3.1. In Table 3.3, we present our model of exploratory search process, which consists of ten features characterizing the exploratory search process. The ten features of the model express typical exploratory search *behaviors* and are mostly related to the exploratory search characteristics.

We did the evaluation of the model in two times, and the preliminary evaluation led to an extended model incorporating the notion of transitions between features; this is this model which we used to design the methods reported in Chapters 4 and 5. In the following Sections 3.4 and 3.5 we describe the first evaluation of the model, which leads us to extend it, and a second assessment of the extended model.

3.4 Preliminary evaluation of the model of exploratory search

Exploratory search is a particular information seeking activity. Our model of exploratory search process satisfies all the principal steps of an information seeking process described in [35]. We associated the features of our model to Marchionini

A. Define the search space (char. 9, 10): The user starts her search session with an anomalous state of knowledge as a general context of search. She has a lack of knowledge and a vague objective of search, and no specific plan to attain it. She will find an approach to her problem and may find an angle of attack.

B. (Re)formulate the query (char. 11): The user (re)formulates the problem with a fluctuating uncertainty. It can be an explicit or implicit formulation: depending on the user interface, the user may use the search bar or keep in mind her query.

C. Gather information (char. 5, 6): The user might not have one precise answer but an aggregate of relevant information which will help her go further in her reflection and exploratory search process.

D. Put some information aside (char. 3, 4, 7): Throughout the search session, the user might put some information aside. She will probably come back to it to pursue/re-start the exploration later.

E. Pinpoint result(s) (char. 2): The user wants more information on one element (query, answer, or the link between them, etc.). This feature is related to sense-making activities such as verifying information.

F. Change goal(s) (char. 1, 4): The user will change or specify the objective/goal of search

G. Proceed backward/forward (char 1): The user can accomplish backward or forward steps when the pathway followed is not suitable for her.

H. Browse results: The user browses or scans the results given by the system.

I. Analyze results (char. 10): The user selects one or multiple filters (or facets) to explore the information space. She will try to fit the results into an analysis framework (relevance of the results). Then, she will identify and analyze all results and possible paths that can be relevant.

J. Stop the search session (char. 8): The user might never end her exploratory search. She can stop it for multiple reasons, and she may continue the search a few hours/days/weeks/months/years/... later.

Tab. 3.3: Initial model: the ten features of exploratory search
The features related to one or several characteristics of exploratory search is indicated with "*char X*"

and White's description of information seeking model, and we can clearly observe that the two models match:

- Recognize a need for information (*Feature A*)
- Accept the challenge to take action to fulfill the need (*Feature A*)
- Formulate the problem (*Feature A*)
- Express the information need in a search system (*Feature B*)
- Examine the results (*Features E, G, H, I*)
- Reformulate the problem and its expression (*Feature B and F*)
- Use the results (*Features C, D and I*)

From this mapping we can conclude that our model of exploratory search is indeed an information seeking model.

In this section we want to verify the relevance of the model. We compare it to the actual exploratory search behaviors of three information seekers. In this evaluation we want to evaluate the following hypothesis:
H1: in the user's exploratory search sessions, we only identify the model's features.

3.4.1 Protocol

The relevance of the model was tested by comparing it to the actual behaviors of three information-seekers performing an exploratory search task. We ask three users to perform an exploratory search on the exploratory search system Discovery Hub[1] (see a. in Table 3.6).

3.4.1.1 The test's protocol

The evaluations follow this protocol:

- **Discovery step of the exploratory search system** (the day before the exploratory search session)
  1. Presentation of the test's goal
  2. Discovery scenario: interactive demo presenting the exploratory search system and all its features (see A. in Table 3.7). Thanks to that, users learned how to use the system and its different features (filters features,

[1] http://discoveryhub.co/

explanation features...), and they did not waste time to discover the system when they will perform the search the day after

- **Exploratory search session**
  1. A reminder of the test's goal
  2. A reminder of the exploratory search system's features
  3. Definition of a personalized search task (the query)
  4. Test session: The participants perform their exploratory search. Each search session was recorded with a screen recorder[2]
  5. Debriefing session: the participants watch the video of their search session and commented their choices, their actions, their thoughts, etc. These explanations were again recorded with another screen record.
  6. End of the test and acknowledgment

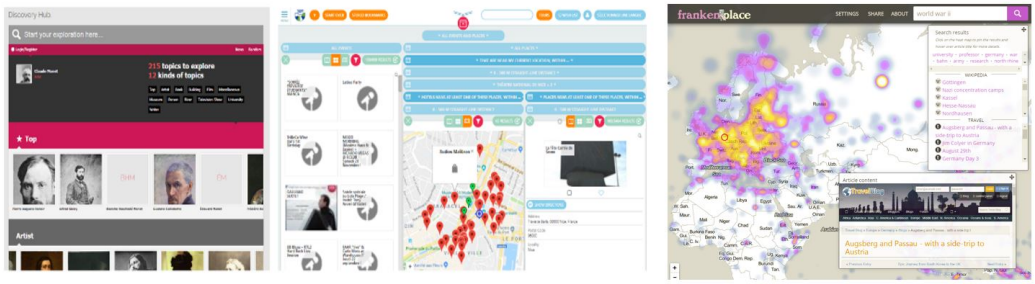
All the exploratory search sessions are about twenty minutes. We supposed that twenty minutes will be enough to show a large sample of exploratory search behaviors.

3.4.1.2 The elaboration of exploratory search tasks

In this protocol, we allow the user to choose the search task and the query. Indeed, users' engagement is really important in user tests, especially in the test of exploratory search systems. It is this engagement, and the user' motivation to explore a subject, which we wanted to induce in this test. We are talking about *ecological tasks*. In other words, the task should suit the user's profession and/or interest to be realistic. For example, "*write an article on topic X*" is right for journalists or bloggers because it is in line with their daily activity and profession.

Then, in these evaluations, we ask to the participants their personal interests in order to propose personalized exploratory search tasks. In the tests, the tasks are: "Discover new board games" for the first participant; "Learn new information about the history of free-jazz" for the second one, and "Learn new information about Senegal" for the third one.

[2] We used Silverback, (`https://silverbackapp.com/`) and Apowersoft (`https://www.apowersoft.com/free-online-screen-recorder`)



(a) (b) (c)

Fig. 3.6: The three different interfaces of the exploratory search system used in the model's evaluation: (a) Discovery Hub, (b) 3cixty, and (c) Frankenplace.

3.4.2 Preliminary results

In the exploratory search records analysis of the three records, we want to verify whether our model reflects the reality of user's exploration. In this section we present the records analysis and the first results.

3.4.2.1 Records analysis

In this analysis, we used the recorded videos together with the comments. The analysis consists in verifying the presence of the model's features in the users' exploratory search activity. The analysis was performed by the designer of the model, by checking the presence of the exploratory search features in the videos and the users' comments, using indicators of this presence. For example when the user is scanning the result list and says "I briefly explore the results list, just to have an idea of the retrieved results", the identified feature is *H. Browsing results*.

Following this methodology, for each video, we write down the different chains of the different model's features users used in their exploration. Table 3.6 shows an example of one exploratory search session analysis on Discovery Hub.

3.4.2.2 Results of the analysis

In this analysis of the three records on Discovery Hub, we do find the features of our model of exploratory search in users' pathway. Table 3.4 lists the number of identified model's features for each exploratory search session. It shows that all the features are not performed in one single exploratory search session (e.g. Feature D for the first participant). All these elements valid our hypothesis H1: we only find

the model's features in an exploratory search session. We concluded that, even if further evaluation was needed, the first records analysis validate in the first instance the model of exploratory search.

We also find an unexpected result: based on the analysis of the chains of the model's feature, we present a non exhaustive list of the possible transitions between the model's features in Table 3.5. In this table, the left column presents the observed features that appear before the features in the central column. One will notice that a search session always starts with **A.** *Define the search space* and ends with **J.** *Stop the search session*. The right column presents the features that can appears after the features in the central column. The transitions list was an unexpected result but turned out to be extremely relevant to better understand the exploratory search process. These results led us to complete our model with the **notion of possible transitions between exploratory search features**.

3.4.3 The first version of the extended model

These first results allowed us to valid and extend the model. Indeed, as explained previously, the transitions between the model's features give a really interesting information about the exploratory search process. These transitions express exploratory search behaviors that all exploratory search systems should facilitate.The list of these possible transitions is not exhaustive because further evaluation may identify new ones.

In our model of exploratory search, as well as in Ellis' model, there is no unique order between the features. The user follows her own search session pathway, according to her thoughts, her expertise in the field explored, the elements of information she encounters, and so on. A user might not perform all of them in an exploratory search session, as shown in Table 3.4. Therefore, different orders are followed by the users when performing their exploratory search session, and it is very informative to identify the transitions that exist between the features: they reflect the important interaction and interdependence between the user and the system in an exploratory search.

Tab. 3.4: List of the model's features identified in the exploratory search session records

Exploratory search systems	Participant	A	B	C	D	E	F	G	H	I	J
Discovery Hub	1	1	10 *(including 2 fails)*	6	0	22 *(including 1 fail)*	8	15	27	9	1
Discovery Hub	2	1	3	3	0	19	2	19	21	3	1
Discovery Hub	3	1	22 *(including 2 fails)*	7	1	45	19	12	29	3	1
Discovery Hub	4	1	2	1	0	14	0	12	9	6	1
Frankenplace	5	1	2	1	0	6	1	7	6	5	1
Frankenplace	6	1	5	1	0	13 *(including 1 fail)*	2	10	12	2	1

Continued on next page

Tab. 3.4 – *Continued from previous page*

Exploratory search systems	**Participant**	**A**	**B**	**C**	**D**	**E**	**F**	**G**	**H**	**I**	**J**
Frankenplace	7	1	1	4	0	5	0	3	3	0	1
Frankenplace	8	1	2	1	0	13 *(including 2 fails)*	4	3	11	7	1
3cixty	9	1	11 *(including 3 fails)*	0	2	7	0	2	10	6	1
3cixty	10	1	6	1	2	5 *(including 2 fails)*	3	6	11 *(including 2 fails)*	2	1
3cixty	11	1	5 *(including 1 fail)*	1	2	3	7	9	11	1	1

Continued on next page

Tab. 3.4 – *Continued from previous page*

Exploratory search systems	**Participant**	**A**	**B**	**C**	**D**	**E**	**F**	**G**	**H**	**I**	**J**
3cixty	12	1	8 *(including 2 fails)*	17	19	16	3	5	40	36	1

Tab. 3.4: List of the model's features identified in the exploratory search session records

In Section 4 and 5, we based our evaluation and design methods on this version of the extended model of exploratory search with the possible transition list described in Table 3.5. In this version of the possible transitions, the two transitions in bold with asterisks were not observed in the first records analysis but were inferred. We added them because they seemed natural to us. We give here some examples:
G $\rightarrow$ **F**: When the user does backward or forward steps, she can have an idea and change her goal(s) of search.
I $\rightarrow$ **G**: When the user analyzes the relevance of a result, she also analyzes the pathway that took her there. If she is not satisfied, she can do backward or forward steps to restore a satisfactory state.

Marchionini in [34] presents a model with three types transitions between the sub-processes: *default, high probability* and *low probability*. At our stage of knowledge on the possible transitions we cannot assume that there is different probabilities between the transitions of our model's features. Moreover, the frequency of occurrence presented in Table 3.4 does not allow to establish probabilities of appearance because the transitions depend on the user, the task, the system's interface (some transitions are not possible in all exploratory search systems, e.g no possibility to bookmark a result from the result list). Then, in Table 3.5, we want only to list all the observed transitions the users wanted to do in their exploratory search session, regardless of the systems.

3.5 Second evaluation of the model of exploratory search

Additional evaluations of the model, following the same protocol, were realized in parallel with the design and evaluation of the two methods described in Sections 4 and 5.

3.5.1 Protocol

Nine exploratory search sessions were performed on three exploratory search systems chosen for their very different user interfaces: one on Discovery Hub[3], four on 3cixty[4] and four on Frankenplace[5] (see b. and c. in Figure 3.6).

[3] http://discoveryhub.co/
[4] https://nice.3cixty.com/
[5] http://frankenplace.com/

Previous features	Feature	Next features
NA	A	B ; J
A ; F	B	G ; H ; I ; J
D ; E ; I	C	D ; E ; F ; G ; H ; J
E ; I	D	C ; F ; G ; J
G ; H ; I	E	C ; D ; F ; G ; J
C ; D ; E ; **G*** ; H ; I	F	B ; H ; I ; J
B ; D ; E ; H ; **I***	G	E ; **F*** ; H ; I ; J
B ; F ; G ; I	H	E ; F ; G ; I ; J
B ; F ; G ; H	I	C ; D ; E ; F ; **G*** ; H ; J
all	J	NA

Tab. 3.5: Non exhaustive list of possible transitions in an exploratory search session. The transitions in bold with asterisks were not observed in the analysis of the first records.

Participant #7 (on Frankenplace):

A → B → H → E → C → G → E → C → G → H → E → G → H → E → C → E → C → G → J

Participant #10 (on 3cixty):

A → B → F (failed) → G → H → E (failed) →G → B → H → F → G → H → E (failed) → H → E → I → F → B → H → B → H (failed) → B → H → E → C → D → H → E → I → D → G → H → G → F → B → J

Participant #4 (on Discovery Hub):

A → B → H → I → B → H → E → G → H → E → C → E → E → G → E → I → G → H → E → I → G → E → G → H → E → G → E → I → G → H → E → G → H → E → G → H → F → E → I → G → J

Tab. 3.6: Records analysis: example of correspondence between users' activity and our model features.

In this second evaluation we reduce the time of the search session to ten minutes because the first evaluations showed that in ten minutes the users can do a quite advanced exploratory search. Thus, we think that ten minutes were sufficient for the next evaluations.

3.5.1.1 The test protocol

The used protocol was the same as the one described in Section 3.4. The *Discovery Scenarios* of the tested system used for the exploratory search system's interactive demo are described in Table 3.7.

3.5.1.2 The elaboration of exploratory search tasks

As we said in the preliminary evaluation, where the protocol is the same for all the exploratory search systems, we cannot give one same task for the three systems. In this protocol, we allow the user to choose the search task and the query as often as we can. We only impose a predefined task for 3cixty, but the task was sufficiently vague for a personal interpretation (e.g. it could be interesting places or events such as museum, restaurant, concert...). Thus, the way we define *ecological* exploratory search tasks and queries differs for the three systems:

- **Discovery Hub.** As in the preliminary evaluations, we ask to the participant her personal interests in order to propose a personalized exploratory search task. The participant wanted to "discover Snooker[6]"
- **3cixty.** The system proposed a database only on the city of Nice (south of France), so we propose one task for all the participants, which can concern everybody in this context: "You want make discover the city for your friends, and in order to do that, you want to plan an outing (find an event) in Nice". Note that all the participants live in Nice or in the area, and they know this city.
- **Frankenplace.** The system offers a very unique user experience. Indeed, the results are presented with heat map on a world map (see Figure 3.6). The number of results depends on the database and the subject explored. We ask the participants to choose the most interesting query on the *examples* proposed by the system on its home page. We needed a query that offers a rich amount of result for a richer exploration. Indeed, in the design phase of the protocol, we realized that only few queries offer on this system a large number of results.

[6]Snooker is a cue sport

A. **Discovery Hub**: "A user heard about Claude Monet on the radio. She wants to know more about this impressionism painter and goes on Discovery hub. She reads the tutorial, and the descriptive text below the search bar. Then, she writes "Claude Monet" on the search bar and launches the query without specifying the characteristics. The results are classified in categories and carrousels. First, she wants to clarify her knowledge on Claude Monet and she clicks on its icon. A pop up opens and she reads the little descriptive text of Wikipedia. She goes back to the result list. She is interested in Camille Pissarro and opens the corresponding pop up. She clicks on the explanation features in order to learn why the result is recommended by the system, what is the link between Claude Monet and Camille Pissarro. Then, always on the descriptive pop up of Camille Pissarro, she reads the associated tags and clicks on "Paul Cézanne". A description of Paul Cézanne appears. She returns on Camille Pissarro's description using the breadcrumb at the top of the pop up. She adds Camille Pissarro to her bookmarks and launches a new search on the painter from the button "run an exploration"."

B. **3cixty**: "A user lives in the city of Nice (France). Two of her friends will come a weekend to visit her. She wants to organize the weekend in order to make discover the city and to find them an accommodation. She goes on 3cixty, reads the text describing the system. She clicks on the tutorial button ("?") in order to better understand the interface and what she can do. She clicks on "Explore all events & places" and again on the tutorial button. The user uses the filter features with the "current location" specification and specifies 800 meters in order to get suggestions of hotels, places and restaurants around her apartment. She selects five results and clicks on "make a new box" and analyses them with the other two views: (1) descriptive view, and (2) the map. She adds one result on her wish list on the descriptive view."

C. **Frankenplace**: "After visiting Turkey, a user wants to know more about Alexander the Great and goes on Frankenplace. She writes on the search bar "Alexander the Geat" and launches the query. The results are presented on an heat map and the user observes that they are mainly concentrated on Greece and Turkey. The user zoom in to get refined results and more precision. An overview of the results appears in a box when the user's mouse pointer hovers the Izmir zone. She scans them and clicks on the zone to fix the results list, which allow an analysis of the results box with more precision. Then, she analyzes the associated tags and moves the mouse over the different results from "Wikipedia" and blog articles in the "TRAVEL" section. She clicks on a tag, and continues her exploration on an other map zone."

Tab. 3.7: Evaluation of the model: the discovery scenarios for the three systems

If the user chose a query by herself, we cannot be sure that the system will present enough results for a ten minutes search session. For all these reasons, when one example matches with participant's interest, we ask the participant to launch a query on the subject (e.g. Castle, Bronze Age, and Olympics). Only one participant was not interested in the example and she choose by herself a subject of exploration : Bruce Springsteen.

3.5.2 Results

As in the preliminary evaluations, we want to verify whether our model reflects the actual exploratory search behaviors. In that sense, we still want to evaluate hypothesis *H1: in the user's exploratory search sessions, we only identify the model's features*.

3.5.2.1 Records analysis

Like the preliminary evaluation, the analysis was performed by the designer of the model, who follows the same analysis methodology: she checked the presence of the exploratory search features in the videos and the users' comments, using indicators of this presence. Table 3.6 shows three examples of exploratory search session record analysis on Frankenplace (participant #7), 3cixty (participant #10) and Discovery Hub (participant #4). When a user failed, for any reason, to do one feature, we noted them X (failed).

As the first evaluation, for each of these exploratory search sessions, we do find again the exploratory search features of our model. Moreover, in the twelve exploratory search sessions, we still do not find any other activity that do not correspond to one model's feature. Considering that we do not find any other activity than our model's features, we can say that our model of exploratory search represents exploratory search behaviors involved in an exploratory search task. The hypothesis H1 is validated, and so is our model.

3.5.2.2 New observed possible transitions

We can therefore say that our model can express the users' activity during an exploratory search task. This time, the possible transitions were an expected result, and we list again the observed ones. With the additional records analysis, we extended Table 3.5 with new observed possible transitions, exposed in Table 3.8.

Previous features	Feature	Next features
NA	A	B ; J
A ; **B** ; F ; **G** ; **H** ; **I**	B	**B** ; **E** ; G ; H ; I ; J
D ; E ; I	C	D ; E ; F ; G ; H ; **I** ; J
E ; I	D	C ; F ; G ; **H** ; J
B ; **E** ; **F** ; G ; H ; I	E	C ; D ; **E** ; F ; G ; **I** ; J
C ; D ; E ; G; H ; I	F	B ; **E** ; H ; **G** ; I ; J
B ; D ; E ; **F** ; **G** ; H ; I	G	**B** ; E ; F ; **G** ; H ; I ; J
B ; **D** ; F ; G ; I	H	**B** ; E ; F ; G ; I ; J
B ; **C** ; **E** ; F ; G ; H	I	**B** ; C ; D ; E ; F ; G ; H ; J
all	J	NA

Tab. 3.8: Extented version of the non exhaustive list of possible transitions in an exploratory search session. The transitions in bold were non observed in the precedent version of the table (Table 3.5).

Note that the new table is still non exhaustive and, as our model of exploratory search, it can evolve with further records analysis.

The inferred possible transitions, with in bold with an astreriks in Table 3.5 were observed in the second evaluation, in the analysis of further exploratory searches. Table 3.8 presents an extended version of the observed possible transitions. These are two versions of a non exhaustive list of the possible transitions between the features of our model. This list can be completed with further analysis of exploratory search sessions.

3.5.3 The extended model

With the additional records analysis of the second evaluation, we validated the model but also discover new possible transitions between the model's features. These new elements led us to extend our model's features with the new observed possible transitions, exposed in Table 3.8. The new version of the extended model is based on the model's features (Table 3.3) and on the new version of the transitions (Table 3.8). Note that the new table is still non exhaustive and, as our model of exploratory search, it can evolve with further records analysis.

3.5.4 Discussion

In this second analysis we observe that, sometimes, depending on the exploratory search system, some features cover several actions. For example on Frankenplace, H. *Browsing History* refers to scan the heat map, and scan the associated results and tags on the window on the right side of the screen; on Discovery hub, it refers to scroll the result list and the carousels; and finally, on 3cixty, it refers to scroll the result list, or observe the results on a map. We voluntarily choose a model of the exploratory search process that can be used on all exploratory search systems. It means that a model's feature can cover several actions depending on the proposed interface, but describes one exploratory search activity. The designer or the evaluator of the system has to specify in her analysis which actions correspond to which feature.

We also observe that sometimes, users failed to do the feature they wanted. For example, in Table 3.6, the participant #10 failed at four times. In these cases, there are two observed reasons: the users do not know how to use the system, or how to achieve an action (which we associate with the corresponding feature in the analysis). We are talking about usability issues (when a user wants to do an action, which is possible, but fails because she does not know how to do it) and missing system's features (when a user wants to do an action, but the system does not offer the possibility to do it; e.g. if a user wants to bookmark a result and the systems does not offer this possibility). Actually, the failed features are really informative and give clues to improve the system in term of usability and proposed features. Indeed, these failed features show how the system should support the exploratory search behaviors.

In this analysis, we observe that, rarely, some model's features iterate. We explain here the observed situations:

- **B** $\rightarrow$ **B**: when the user refines the query (e.g. on 3cixty, the users can refine the launched query by using filters features).
- **E** $\rightarrow$ **E**: when a user discovers a result and clicks on an associated tag for discovering the tag's description (e.g. on Discovery Hub tags are associated to result's description).
- **G** $\rightarrow$ **G**: when a user performs several backward steps.

From this observation, we can suppose that the possible transitions we described in Table 3.8 are not the same for every system, according to their interface (e.g. there is no possibility to do the transition $E \rightarrow E$ on 3cixty).

Finally, we observe some *patterns* in the users' behaviors. Here, the term *pattern* refers to identical series of features. Further analysis of exploratory search behaviors may show more patterns. We think that strategies of exploratory search can appear from these patterns. This may be really informative for a better understanding of exploratory search process. Indeed, strategies are exploratory search behaviors that every exploratory search system also should support. We give here some examples observed in the analyzed exploratory search sessions:

- **Start the search session:** all exploratory search sessions start with *A. Define the search space* → *B. Formulate the query* → *H. Browsing results*, which means that when a user defines the goal of search (A), she formulates the query (B) and scans the result list (H).
- **Reformulate a query:** all the query reformulations are identified with *F. Change goal(s)* → *B. Reformulate the query*. In other words, before reformulating the query (B), the user may change her goal of search (F).
- **Bookmark results:** all the bookmarked results are preceded by *I. Analyze results* → *C. Gather information* → *D. Put some information aside*. It means that the user first analyzes the relevance of a result (I), gathers information that she finds interesting (C) and puts it aside as a bookmark (D) for individual reason.

3.6 Conclusion

In this section we wanted to find or elaborate an *approximate* model of exploratory search which will serves as a basis to our evaluation methods.

In this design process, we first established the eleven characteristics of exploratory search. We used them as an analysis grid to evaluate five candidate models of information seeking. Our goal was to establish if we can choose without modification one of them. None of the models checked all the characteristics of exploratory search and we had to adapt one of them to the exploratory search process. The selected model is Ellis' model because its description of the search process is non-linear, and without predefined sequences. Moreover, the author describes explicitly a model that can be used for the design of recommendations.

Thus, we designed our model of exploratory search on Ellis' model and we evaluated its relevance by verifying if it effectively reflects the activity of three successive groups of participants using respectively one of the following exploratory search systems: Discovery Hub, 3cixty and Frankenplace.

The results analysis shows that in the twelve analyzed records, we can find all the model's features and no other kind of search activity. The model of exploratory search is validated and then expresses the users' activity during an exploratory search task. From the evaluations, we also find an unexpected result: an non-exhaustive list of the possible transitions between the model's features. These transitions are extremely interesting for a better understanding of users' exploratory behaviors. We extended the model by integrating in it this notion of possible transitions.

The evaluation shows also the importance of a good usability. Indeed, problems of usability are more than a half of the cause of failed features. It shows also what kind of information or interactions the user may wants to get a satisfactory exploratory search experience.

In the next chapter we will present the first evaluation method we designed based on our exploratory search model. This method is an inspection method of exploratory search systems' interfaces.

4 An inspection method without users

4.1 Introduction: the need for a simple and inexpensive inspection method

In Chapter 3, we introduce a model of exploratory search. We designed this model with the intent to use it as a basis for the elaboration of evaluation methods of exploratory search systems that effectively assess their capacity of support the exploratory search process. After the model design, we decided to propose two different user-centered evaluation methods for a complete user-centered evaluation of exploratory search systems. These two methods fall into two categories which are commonly used for the users interfaces evaluation: inspection methods and empirical methods. The inspection methods (e.g. *cognitive walk-through, heuristic evaluation, feature inspection, Standards inspection...*) differ from empirical methods (e.g. user test, interviews...) in that way they do not involve users in the evaluation. The inspection of the interface is realized by evaluators, mainly experts in user-centered evaluation (e.g. user experience, usability, persuasive technologies...). These methods are also qualified as *discount*, *i.e.*, low-cost methods [16, 46]. Indeed, interface testing with users requires a lot of time (for the preparation, the recruitment of users, the tests, the analysis...) and often a budget for the recruitment. On the contrary, inspection methods are easier to apply and *save* users for tests at a later stage. Empirical evaluation and inspection methods are not enough by themselves: they are complementary to the other methods. Reporting on usability inspection methods in [47], Nielsen claimed that these methods "*are able to find many usability problems that are overlooked by user testing but that user testing also finds some problems that are overlooked by inspection, meaning that the best results can often be achieved by combining several methods*".

We decide to propose two kinds of model-based methods: an inspection method that can be use all along the design process of the exploratory search system, and an empirical method that involve end-users during the assessment phase, which imply at least a functional prototype (see Chapter 5). We focus in this chapter on the way we designed our *heuristics* of exploratory search and evaluate them.

The work exposed in this section was published in [55].

4.2 Heuristics evaluation as the type of expected inspection method

The evaluation methods we design target the designers of these systems. It means that the computer scientists that will use the methods are not familiar with evaluating their systems in a user-centered way. This include the use of inspection and empirical methods. This fact imply a constrain: we have to design evaluation methods that can be used as they are, without requiring any knowledge about them.

We are focusing in this section on the design of an inspection method which will be based on our model of exploratory search. In the literature, we can find different kind of inspection method. [47] offer a description list of inspection methods (heuristic evaluation, cognitive walk-through, feature inspection...) which can be used in the usability analysis. The heuristic evaluation and cognitive walk-through are the most used ones:

- **Heuristic evaluation:** from all the inspection methods, heuristic evaluation is the "most informal" one [47]. It is an easy method to perform, and can be used several times along the design process, from the first sketches of the user interface to the final product. Heuristic evaluation is therefore compatible with classical development life cycles, as well as prototyping and agile approaches. A heuristics-based evaluation method allows the identification of issues and their correction in a specific domain. In such a method, evaluators use a set of established principles, called *heuristics* or *criteria*, to inspect the user interface. For each element of the interface, they will determine its compliance with the appropriate principle/heuristic. Several expert evaluators (generally between 3-5 evaluators) evaluate the interface individually. They will combine the inspection reports and confront what they identified. Nielsen's heuritics [45] and Bastien and Scapin's ergonomic criteria [5, 65] are a good examples of heuristic evaluation in usability domain.
- **Cognitive walk-through:** this inspection method is a task specific approach, in contrast with the heuristic evaluation which is more *holistic* in the issues identification[1]. Indeed, Cognitive walk-through method is usually used to evaluate the learnability of software interfaces [68]. As heuristic evaluation, this method can be applied at an early stage in the design process of the evaluated system. In cognitive walk-through the evaluators define the task or

[1] From: https://www.interaction-design.org/literature/article/how-to-conduct-a-cognitive-walkthrough

tasks system's users would do. The idea is to verify if any action leads to the next correct action to achieve the identified task.

Cognitive walk-through presupposes a good knowledge on users' tasks, in our case exploratory search tasks. However, we cannot predict what actions are needed to achieve one feature. Indeed, each interface of exploratory search system is different. In addition, we cannot predict the transitions between our model's features (see Tables 3.5 or 3.8), all the more as we have not ourselves an exhaustive knowledge of transitions. It would be quite difficult to design such an inspection method. Thus, we choose to design a heuristic evaluation method because it matches to our needs: it is an inspection method with a simple protocol and which can be easily used by non-experts. In addition, the principles (or heuristics/criteria) can be derived from our model of exploratory search. These principle can be fairly generic and applied on all kind of exploratory search systems. Indeed, these principles will allow to verify whether the system effectively support model's features and the transitions between the features.

4.3 The model-based design process of our heuristics of exploratory search

The inspection method we want is divided into two elements: the heuristics of exploratory search and a procedure to use them (see Section 4.4). In this section we focus on heuristics and explain that they are based on our model of exploratory search process (in its first extended version) we designed and presented in Chapter 3. The heuristics must help assess that the evaluated exploratory search system support the model's features and facilitate the transitions depicted in Table 3.3 and 3.5. Indeed, the evaluation method must take into account the transitions, in order to make the exploratory search process easier. In other words, the heuristics translate the features and the transitions into concrete interface elements or possible actions required to effectively support exploratory search behaviors. Designed appropriately, these elements will help the user perform her exploratory search task.

In this section, we present the design method of our heuristics of exploratory search and their two different forms.

4.3.1 Heuristics' design methodology

The heuristics of exploratory search are based on the model's features and transitions, in the sens that they translate them in concretely elements or possible actions

required to support exploratory search behaviors. However, in the design process of the heuristics, we realized that one heuristic is not necessarily linked to only one feature or to only one transition. So we had to cluster features and transitions.

We fist start with the transitions between the model's features: we listed the different ways to facilitate them. We chose to start with the transitions because they are more linked to concrete system's features than the model's features (e.g. transition E,I $\rightarrow$ D is related to *bookmarks* feature). In addition, we saw in Section 3.4 what kind of systems' features support these transitions. Thus, we started the heuristics design from a baseline, with concrete elements (e.g. necessary system's features) observed in the exploratory search sessions in Sections 3.4 and 3.5. For example, F (*Change goal(s)*) $\rightarrow$ B (*(Re)Formulate the query*) can be realized in different ways and implies different needs in terms of interface elements: the user must be able to change easily her search at any time during her search session. This includes different situations, and that is why two heuristics are proposed:

- The search bar or the bookmarks are accessible anytime during the search session.
- The system should allow to start a new search from the search bar but also from an element of the interface (e.g. a result, a world/element on the result description...).

After that, we cluster the transitions corresponding to the same action with the same interface elements. For example, we cannot cluster the previous heuristics, but the transitions E (*Pinpoint result*) $\rightarrow$ D (*Put some information aside*) and I (*Analyze results*)$\rightarrow$ D can both be facilitated by using bookmarks.

Notice that transitions C, D, E, G, H, I $\rightarrow$ F does not have linked heuristics because there are no interface elements that facilitate the transition to F (*Change goal(s)*). Indeed, feature F cannot be translated by a physical action: it occurs when the user decides to change her search goal.

We then add heuristics that facilitate model's features, and we also based the heuristics we propose on the observed exploratory search sessions of Sections 3.4 and 3.5. We note the elements of interface on the four systems analyzed that allow or facilitate the performance of a feature (e.g. a search bar for feature B). We also propose solutions for failed model's features related to a missing system's feature (e.g. an browsing history). Same heuristics can refer to different features, for example we cluster "A & B" and "E & I" sometimes more than once.

At this time we have our list of heuristics of exploratory search which support the exploratory search behaviors.

Furthermore, we notice that the heuristics can be classified into different screens: home page, list of results/data, elements' description, browsing history/breadcrumb, and *any screen*. It appears to us that this classification would facilitate the usability of the heuristics for computer scientists. And to this end, sometimes, we need to split some heuristics in two, e.g. the previous transitions E,I $\rightarrow$ D:

- **List of result/data:** The system should allow the user to bookmark elements from the result list.
- **Elements' description:** The system should allow the user to bookmark elements from the elements' description.

The evaluator will be able to use these classified heuristics without understanding or knowing the model of exploratory search process or the transitions between the features to evaluate the exploratory search system. She will only use the heuristics and verify if the system satisfies all of them.

All the heuristics provide elements that support at least one feature or transition. Thus, if at the end of an evaluation of a given exploratory search system we find that the system satisfies all the heuristics,we can conclude that it supports the exploratory search process and behaviors. As mentioned earlier, this evaluation of exploratory search process should be performed by persons familiar with the exploratory search process (e.g. exploratory search systems designer, UX designer...). The level of knowledge on exploratory search can vary but a certain level of experience is required: the evaluators must be familiar with the project and the exploratory search concept. Indeed, they will use the heuristics as well as their own experience in exploratory searches to analyze and improve their exploratory search system.

4.3.2 The heuristics of exploratory search: the classic form

Here we introduce a list of exploratory search heuristics based on our model of exploratory search. For an easy use, they are classified according to the user interface screens (e.g. home page, list of results...). The evaluator can inspect her system by following our heuristics' presentation.

Here we present the twenty two heuristics of exploratory search we designed. They are classified into five different screens: the *home page*, the *list of results/data*, the *elements' description* (referring to the results' description and/or the query's description), the *browsing history/breadcrumb* and the heuristics which can be applied on *any screen*.

1. **On the home page**

- **Description of provided information:** from the very start, the user should understand what kind of information she will find by using the system (the explored field, the mediatype...). She may know what kind of query or search she can do. (A)
- **Tool for defining the search space:** the system use tools to help the user to define her search space (e.g. query auto-completion feature, suggestions, automatic correction or filtering tools, etc.). (A & B)
- **Proximity of the search bar/system's description:** the search bar (or bookmarks, filtering tools, etc.) and the system's description are always nearby. (Transition A $\rightarrow$ B)

2. **On the list of results/data**
   - **List of results after query formulation:** the system should present the results list after a query (re)formulation and a change of goal(s). (Transitions B, F $\rightarrow$ H)
   - **Picture/snippet associated to a result title:** the system should associate to the title of the result a picture and, if it is possible, a result snippet. The idea is helping the user to understand and evaluate the results list faster. (I)
   - **Result presented WRT result type/main activity:** the system should present the result/data in a relevant way according the type of the result/data and the main activity (list, graph, carrousel, parallel interfaces...). (I)
   - **Bookmarking from result list:** the system should allow the user to bookmark elements from the result list. (Transitions E, I $\rightarrow$ D)
   - **Way to access to element description:** the system should show in the results list the way to access to the elements' description (e.g. more details button). (Transitions H, I $\rightarrow$ E)
   - **Most relevant results displayed first:** the system should first display the most relevant results (e.g. a top selection list, the similarity or even the distance between the query and the result, etc.). (Transition B $\rightarrow$ I)
   - **Filters features available:** the system should make available filters features that help the user to understand and analyse the results' list. These filters can be proposed by the system and/or chosen by the user (e.g. common facets). (Transition F $\rightarrow$ I)
   - **Explanation of the results:** the system should explain why a result is proposed from the results' list. This explanation should help the user to analyse its relevance. (Transition H $\rightarrow$ I)

3. **On the elements' description**

- **Query description:** if there is query formulation, the system should provide a description of the query. Sometimes the user may explore unknown subject and need to learn more about it. (E & I)
- **Several types of information for one result:** the system should give enough information to provide a better understanding of the result. Every result should propose a description (text, picture(s), video, maps, sources...), the associated tags, characteristics, links, etc. These information enable the user to analyse the relevance of the result. The user mustn't need to go elsewhere (another search engine or website) to search the explanation. (E & I)
- **Explanation of the results:** the system should explain why the result or element is proposed (why the user would find it relevant). Indeed, the link between two elements should be described (e.g. the link between the query and the result: the characteristics in common, a chosen filter...). (E & I)
- **Comparison features:** the system should offer comparison features between at least two elements (including the query, different results...). This point is really important, especially if the user is often led to compare several element of information in her goal task or exploration. (E & I)
- **Bookmarking from the description:** the system should allow the user to bookmark elements from the elements' description. (Transitions E, I $\rightarrow$ D)

4. **On browsing history/breadcrumb**
   - **Element description already read:** in the browsing history or the breadcrumb, the system should show the elements' descriptions the user already read. (Transitions G $\rightarrow$ E, and G $\rightarrow$ I)
   - **Presence of a shortcut to the results lists:** in the browsing history or the breadcrumb, the system should show where the user was browsing the results' list. (Transitions G $\rightarrow$ H, and G $\rightarrow$ I)
5. **On any screen**
   - **Always accessible search bar and bookmarks:** the search bar or the bookmarks are accessible anytime during the search session. (Transition F $\rightarrow$ B)
   - **New search launchable from search bar/any element of the interface:** the system should allow to launch a new search from the search bar but also from an element of the interface (e.g. a result, a world/element on the result description, etc.). (Transition F $\rightarrow$ B)

- **Easily or always accessible browsing history/breadcrumb:** the system should propose an easily (or always) accessible browsing history breadcrumb and back/next buttons. (Transitions B, D, E, H, I $\rightarrow$ G)
- **Launched query always accessible:** the system should help the user to keep in mind her goals of search (e.g. the user can see, at any time, the query and the exploratory search pathway throughout breadcrumb or browsing history). (G)

4.4 The protocol

In Section 4.3.2 we present our heuristics of exploratory search. In this section, we present two other adapted version of the heuristics for an easy use. We also propose a simplified evaluation procedure to use them.

4.4.1 Usability of the heuristics of exploratory search

In a inspection evaluation, the evaluators can print them use them as they are. Nevertheless, we thought about ways in which the usability of the heuristics would be easier. We propose here two other version of the *classics* heuristics of exploratory search: an interrogative form version, and a online version of this interrogative form. Indeed, the first computer scientist's feedbacks on the classic version of the heuristics reveals that heuristics in an interrogative form were more *understandable* for them. As we said earlier, they are not familiar with *inspection methods*, and an interrogative form allow a better understanding on what they have to do in such evaluation (e.g. seek on the interface if a specified element is effectively present).

In sum, the heuristics can be used:

- In a paper form for the two version of the heuristics: *classic* and *interrogative*;
- Online with a Google Form which use the *interrogative* form of the heuristics.

4.4.1.1 Heuristics of exploratory search: the form format

Here we present the heuristics of exploratory search in their interrogatory form. As the classic form, they are classified into five different screens: the *home page*, the *list of results/data*, the *elements' description*, the *browsing history/breadcrumb* and the heuristics which can be applied on *any screen*.

1. **On the home page**

   Q1.a. Is there a text describing the system and/or a tutorial (or other)?
   From the very start, the user should understand what kind of information she will find by using the system (the explored field, the mediatype...). She must understand from the starting screen what kind of query or search she can perform.

   Q1.b. Does the system propose features or tools that help the user to define the search space?
   The system use tools to help the user to define her search space such as: query auto completion feature, query suggestion, automatic correction, filtering tool, other...

   Q1.c. Are the search bar (or bookmarks, filtering tools, etc.) and the system's description nearby on the home page?

2. **On the list of results/data**

   Q2.a. After formulating a query (or refining one), does the system present the result list?

   Q2.b. On the result list, does the system associate to the retrieved results several information about them?
   The idea is helping the user to understand and evaluate the results list faster (e.g. a title, a picture, a result snippet, other...)

   Q2.c. Does the system present the results or data in a suitable way according to the main activity (e.g. list, graph, carrousel, parallel interfaces...)?
   For example, if the main activity is to compare elements, a parallel interface would be more appropriate. On the other hand, if the main activity is to explore a large amount of data, a graph visualization would be more relevant.

   Q2.d. Can the user bookmark elements from the result list?

   Q2.e. Does the system indicate the way to access to the elements' description from the results list (e.g. more detail button, other...) ?

   Q2.f. On the results list, does the system first display the most relevant results?
   e.g. top selection list, Features showing the degree of similarity, Features showing the degree of distance between the query and the result, other...

   Q2.g. Are there filters features that help the user to understand and analyze the results list?
   These filters can be proposed by the system and/or chosen by the user (e.g. common facets).

Q2.h. Does the system explain why a result is included in the results list?
This explanation should help the user analyze its relevance (e.g. explanations features, other. . .)

3. **On the elements' description**
Q3.a. When a query is launched, does the system propose a description of the query?
Sometimes the user may explore unknown subject and need to learn more about it. A description of the query will help the user to explore and evaluate the results.
Q3.b. Does the system associate in the description of a result enough information to perform an exploratory search?
The system should give enough information to provide a better understanding of the result. Every result should propose a description (text, picture(s), video, maps, sources, other. . .), the associated tags, characteristics, links, etc. These information enable the user to analyze the relevance of the result. The user must not need to go elsewhere (another search engine or website) to search the explanation.
Q3.c. Does the system provide in an element's description an explanation about the link between two elements (query - result, or between two results)?
The system should explain why the result or element is proposed (why the user would find it relevant). Indeed, the link between two elements should be described (e.g. the link between the query and the result: the characteristics in common, a chosen filter. . .).
textitQ3.d. Does the system provide comparison features between at least two elements (including the query, different results, etc.)?
This point is really important, especially if the user is often led to compare several element of information in her goal task or exploration.
Q3.e. Can the user bookmark elements from the elements' description?

4. **On the browsing history/breadcrumb**
Q4.a. Does the system provide a breadcrumb or a browsing history?
A browsing history or a breadcrumb should help the user to make some backward steps when she estimates that the followed pathway is no longer suitable for her.
Q4.b. In the browsing history or the breadcrumb, does the system show the elements' descriptions the user already read?
If your system does not provide any browsing history or breadcrumb, please skip this question.

Q4.c. In the browsing history or the breadcrumb, does the system show where the user was browsing the results' list?
If your system does not provide any browsing history or breadcrumb, please skip this question.

5. **On any screen**
   Q5.a. Does the search bar and the bookmarks are always accessible during the search session?
   Q5.b. Does the system allow the user to launch a new search from the search bar but also in another way (e.g. from an element of the interface)?
   For example it can be from a result snippet, a world/element on the result description, etc.
   Q5.c. Does the system propose an easily (or always) accessible browsing history, breadcrumb and back/next buttons?
   Please skip this question if your system does not provide browsing history, breadcrumb and back/next buttons.
   Q5.d.Does the system show the launched query on any screen?
   The system should help the user keep in mind her goals of search (e.g. the user can see, at any time, the query on the result list and the exploratory search pathway throughout breadcrumb or browsing history).

For a better understanding, we illustrate these heuristics with screenshots when it was possible in a document accessible online: `https://goo.gl/odrS2e`.

4.4.1.2 The evaluation checklist

We propose the heuristics in their interrogative form on an online Google From, we call it the *evaluation checklist*:`http://bit.ly/evaluation-checklist`. We designed this evaluation check list for an easy use of the heuristics: the evaluators can record online their answers for each heuristics. Moreover, the inspection of an interface is performed individually by several evaluators. The online form groups all the evaluators' answers in one unique report and therefore ease the comparison. We suggest the evaluators to perform their inspection with the evaluation checklist, and that is why we tested them in the evaluation of the heuristics in Section 4.5.

4.4.2 Prepare and perform the evaluation

Accompanying the heuristics, we propose a simplified procedure to use it, i.e., a procedure for performing a heuristic evaluation based on our evaluation checklist. This procedure is inspired from the classical heuristic evaluation procedures (see, e.g.,

Nielsen [47]). The procedure consists of three steps: (1) preparing the evaluation; (2) performing the evaluation; (3) analyzing the evaluation results and suggesting recommendations.

1. **Preparing the evaluation**
   - **Forming the evaluation team.** The evaluation is supposed to be directed by an evaluation director, who first recruits two to five evaluators to ensure that a large part of the critical interface aspects (negative and positive) will be identified. The evaluators must be familiar with the exploratory search process and belong to the same design team (computer scientists, UX designers, etc.). Their knowledge about exploratory search can vary but a certain level of experience is required. Indeed, they should understand users' needs and behaviors in an exploratory search session. Experts will notice more interface issues than novices.
   - **Briefing the evaluators.** The evaluators are informed of the goal of the evaluation and familiarized with the system to be evaluated. They are provided with the heuristics in their form format[2] i.e. phrased in an interrogative form and classified into different screens (home page, results list, etc.) in order to facilitate the evaluation process.
2. **Performing the evaluation**
   - **Inspecting the interface.** The evaluators perform the evaluation individually. Using the evaluation checklist, they analyze the interface screen by screen. On a screenshot, they identify for each question the corresponding elements of the interface (e.g. search bar, description the system, result snippet...). They indicate for each element of the interface if it complies the heuristic in its question format (positive aspect - "Yes") or not (negative aspect - "No"). It is very important to record the positive aspects also in order to keep them as is thereafter. For example, consider the evaluation of Aemoo a Semantic Web application supporting knowledge exploration on the Web. When considering the Aemoo home page (see Figure 4.1), this evaluator may notice that it fulfills the heuristic: HOME PAGE > The search bar and the system's description are always nearby.
   - **Reporting the results of the evaluation session.** To document the positive and negative aspects they observed, the evaluators write a personal report in which they include the screenshots corresponding to each observed aspect, and an explanation of why the heuristic is fulfilled or not. Each evaluator sends her report to the other evaluators.
3. **Analyzing the evaluation results and suggesting recommendations**
   - **Group discussion:** the evaluators compare and discuss together their results and see if they missed one or several issues or positive aspects.

[2] http://bit.ly/evaluation-checklist

Fig. 4.1: Illustration of the fulfilled heuristic: "The search bar and the system's description are always nearby". (1) The search bar, and (2) The system's description which is a click away from the search bar.

They discuss how the highlighted issues could be corrected and suggest recommendations (e.g. "add a browsing history" or "add a query auto-completion feature").

- **Reporting the results of the discussion:** The evaluation director and the evaluators write a collective synthesis of the evaluation. This document summarizes the positive and negative aspects with the corresponding heuristics and screenshots, and the resulting recommendations. The report is sent to the rest of the design team in order to take design decisions. If this evaluation procedure is used iteratively throughout the whole design process, it means that each interface improvement will respect the exploratory search process.

4.5 Evaluation of the heuristics of exploratory search

The main goal of our work here is to determine if an exploratory search system supports users' exploratory search behaviors. In order to do that, our heuristics intend to identify system's features, interface elements or possible actions that are required for the proper achievement of the exploratory search task. We assume that an evaluation of the system with our heuristics of exploratory search allows a better identification of these elements. We want to evaluate the following hypotheses:
(H1) users following the evaluation checklist identify more provided or missing

system's features, interface elements or possible actions than users not following the evaluation checklist.
(H2) users following the evaluation checklist perform a more systematic identification than users not following the evaluation checklist.

4.5.1 The method used for the evaluation of the heuristics

We present here the used methodology for the evaluation of the heuristics.

4.5.1.1 The selected participants

In this evaluation we have twenty participants, all of them are computer scientists with any knowledge on exploratory search concept. All participants did not know any of the evaluated systems. Ten participants evaluated Discovery Hub: five with the heuristics (Group A) and five without the heuristics (Group B). The ten other participants evaluated 3cixty: five with the heuristics and five without the heuristics.

4.5.1.2 The material

Exploratory search systems. We asked the participants to evaluate two different exploratory search systems: Discovery Hub and 3cixty. These systems were chosen for their really different interfaces, see Figure 3.6.

Online evaluation checklist. In the users tests we proposed to the participants the heuristics in their form format mentioned in Section 4.4.1.1. More precisely we ask the participants to use the online version: the evaluation checklist[3]. We printed the evaluation checklist because we could not split the screen of the used computer for a parallel presentation of the systel and the online form. We also provide paper and a pen for the users who do not evaluate the systems with the evaluation checklist.

4.5.2 The used procedure in the evaluation of the heuristics

We present here the used procedure for the evaluation of the heuristics of exploratory search for the two groups A and B. From the protocol exposed in Section 4.4.2 we only use and adapt the steps *Briefing the evaluators* and *Inspecting the interface*.

[3] http://bit.ly/evaluation-checklist

4.5.2.1 Scenario

Participants not knowing the evaluated systems at the beginning of the test, we proposed to familiarize them with a scenario-based demo of these systems. The demo was carried out by the evaluator conducting the test. The scenario were also used by Group B in the inspection phase of the evaluation. The scenarios were designed to show all the features and the possibilities of the two systems:

Discovery Hub: "A user heard about Claude Monet on the radio. She wants to know more about this impressionism painter and goes on Discovery hub. She reads the tutorial, and the descriptive text below the search bar. Then, she writes "Claude Monet" on the search bar and launches the query without specifying the characteristics. The results are classified in categories and carousels. First, she wants to clarify her knowledge on Claude Monet and she clicks on its icon. A pop up opens and she reads the little descriptive text of Wikipedia. She goes back to the result list. She is interested in Camille Pissarro and opens the corresponding pop up. She clicks on the explanation features in order to learn why the result is recommended by the system, what is the link between Claude Monet and Camille Pissarro. Then, always on the descriptive pop up of Camille Pissarro, she reads the associated tags and clicks on "Paul Cézanne". A description of Paul Cézanne appears. She returns on Camille Pissarro's description using the breadcrumb at the top of the pop up. She adds Camille Pissarro to her bookmarks and launches a new search on the painter from the button "run an exploration"."

3cixty: "A user lives in the city of Nice (France). Two of her friends will come a weekend to visit her. She wants to organize the weekend in order to make discover the city and to find them an accommodation. She goes on 3cixty, reads the text describing the system. She clicks on the tutorial button ("?") in order to better understand the interface and what she can do. She clicks on "Explore all events & places" and again on the tutorial button. The user uses the filter features with the "current location" specification and specifies 800 meters in order to get suggestions of hotels, places and restaurants around her apartment. She selects five results and clicks on "make a new box" and analyses them with the other two views: (1) descriptive view, and (2) the map. She adds one result on her wish list on the descriptive view."

4.5.2.2 Inspection of the interface

At this point, the procedure differs for the two following groups:

- **Group A (Evaluation with checklist):** the questions of the evaluation checklist, as well as the heuristics of exploratory search, are classified according to the different screens of the exploratory search systems: the home page, the list of results/data, elements' description and browsing history/breadcrumb, and any screen. This classification provides a guidance of the system evaluation that is both easy to understand and use. We asked the participants to follow the form and answer to the questions.
- **Group B (Evaluation without checklist):** after the demo of the system with the scenario, we repeated the scenario and we asked the participants to formulate all the negative and positive aspect that can facilitate or compromise the exploration.

4.5.3 Evaluation metrics

We measured the effectiveness of our heuristic method in terms of precision, recall and F-measure of the participants' answers:

- **Group A (with the evaluation checklist):** each time a participant has to check "Yes" or "No" in a multiple choice list.
- **Group B (without the evaluation checklist):** each time a participant's comment refers to an answer ("Yes" or "No") to the evaluation form.

In [49], Nielsen and Molich evaluate their heuristics by comparing the participants' identified usability problems to a list of problems developed by the authors. We used the same approach here. Then, we created a gold standard by asking two usability experts who possess a good knowledge of exploratory search process to evaluate the two systems with the evaluation checklist. Each expert made the evaluation independently. In a first phase, those two evaluations were confronted in order to evaluate the understandability of the questions. This phase leads to a reformulation of questions 2.a, 3.a and 5.a. Moreover, questions 2.b, 2.f were more detailed for an easy understanding. In a second phase, the two experts re-evaluate the two systems leading to a new computation of their agreement score. All these elements are in Table 4.1

According to Landis and Loch, there is a perfect agreement between the experts after the phase 2. We evaluate the answers analysis using this second gold standard. For

Tested systems	Testing phase	
	Phase 1	Phase 2
Discovery Hub	0,309623431	1
3cixty	0,573275862	1

Tab. 4.1: Cohen's kappa (κ) according to the testing phase and the tested system

Group A's answers analysis: we calculate the number of correctly and incorrectly identified "Yes" and "No" answer. For Group B's answers analysis: we first identified the comments only corresponding to evaluation checklist's answers. Then, we calculate the number of "Yes" or "No" correctly and incorrectly identified. From that, and for each participant, and for each system, we used the following two formulas in order to calculate recall and precision for all Yes, No, and the two answers combined.

$$Recall = \frac{number\ of\ X\ correctly\ identified\ by\ the\ participant(s)}{number\ of\ X\ the\ participant(s)\ should\ have\ identified}$$

$$Precision = \frac{number\ of\ X\ correctly\ identified\ by\ the\ participant(s)}{number\ of\ X\ identified\ by\ the\ participant(s)}$$

Here, *X* refers to provided or missing system's features, interface elements or possible actions

4.5.4 The quantitative and qualitative results

The results were analyzed quantitatively and qualitatively. The qualitative results refers to Group B's answers that are not taken into account in the quantitative results.

4.5.4.1 Quantitative results

As described in the last section, we use precision and recall metrics to analyze the participants' answers. F-measure is the harmonic mean of these two metrics and reflects the test accuracy. Figures 4.2 and 4.3 show the F-measure for the two experiments with the two different systems. In each figure the first bar on the left (green) shows the F-measure for the users following the evaluation checklist (Group A) while the second bar on the right (red) shows the F-measure for the users not

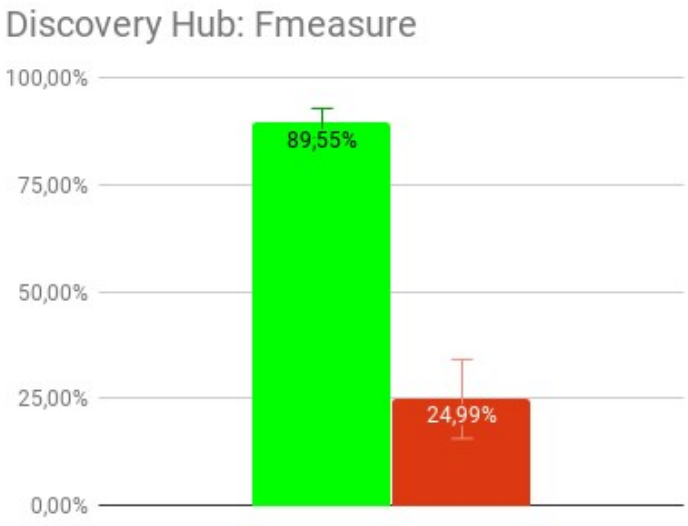


Fig. 4.2: F-measure of the tests on Discovery Hub. The left bar refers to Group A (with the evaluation checklist) and the right one to Group B (without the evaluation checklist).

following the evaluation checklist (Group B). For each bar we also indicated the standard deviation for each group with an error bar.

We can now come back to our two hypotheses:

- For (H1) considering the F-measure of 83.18% vs 14.66% (Figure 4.3) and 89.55% vs 24.99% (Figure 4.2) we can say that the users with the evaluation checklist identify more provided or missing features than users not following the evaluation checklist.
- For (H2), with a standard deviation of 3.35 vs 10.32 and 3.45 vs 9.26, we demonstrate that users following the evaluation checklist perform a more systematic identification than users not following the evaluation checklist.

For both Discovery Hub and 3cixty, the differences between Group A and Group B in terms of Precision and Recall (F-measure) were found to be statistically significant (Mann-Whitney test, $p<0.02$). In other words, the heuristics allow significantly a better exploration-oriented evaluation of the two interfaces.

4.5.4.2 Qualitative answers of Group B

Only participants of Group B provided comments not related to exploration-oriented actions/system's features but related to other aspects such as usability and user experience. We give here a few chosen examples of comments and we link them

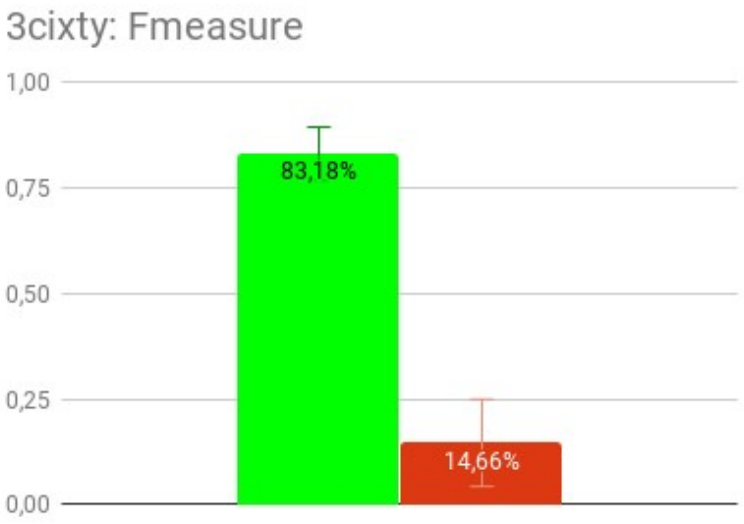


Fig. 4.3: F-measure of the tests on 3cixty. The left bar refers to Group A (with the evaluation checklist) and the right one to Group B (without the evaluation checklist).

to Bastien and Scapin's ergonomic criteria [5]. For Discovery Hub, participant 2 said "On the home page, the text on a picture of grass is difficult to read". This refers to *Legibility criteria*. Participant 3 said about the result list: "I do not know that if I click on the query reminder, it leads to a descriptive pop up of the query". This refers to the *Prompting criteria*. For 3cixty, all the users said that "the result list gives too many results". This refers to the *Information Density criteria*. Then, user 3 proposed "a colour code to help the user distinguish the different kinds of results" (e.g. places, restaurants, hotels), and the four others proposed to "categorize them into predefined categories". These propositions refer to the *Guidance criteria*. Participant 3 added: "if I want to filter the results list I have to click 3 to 4 times, it is too much! The filters feature should be proposed directly on the results list to halve the number of clicks". This refers to the *Minimal actions criteria*.

For the two systems, these types of comments represent an average of 61.4% of the comments. Indeed, Group B's comments about usability and user experience issues represent 47.1% for Discovery Hub vs 75.6% for 3cixty with a standard deviation 14.9 vs 11.5. These results show that the heuristics we propose allow a significantly better identification of provided or missing exploration-oriented features, interface elements or possible actions than without them. Indeed, Group B's participants readily use usability and user experience criteria. The evaluation checklist format of the heuristics provides an efficient framework and guidance for the interface inspection. On the other hand, the guidance can be, in some way, too restrictive in the sense that evaluators may miss other important interface issues related to the usability of the system or the user experience. In fact, the heuristics and their form format are clearly not enough for a complete evaluation of an exploratory search

system. For that purpose, the methods proposed here must be combined to other existing well-known methods on usability and user experience such as Nielsen's usability heuristics [47] or Bastien and Scapin's ergonomic criteria [5].

4.5.5 Discussion

In this evaluation we tested the capacity of the method to help evaluators identify the presence or the absence of system's features, interface elements or possible actions allowing users to develop exploratory search behaviors and to perform exploratory search tasks. The results demonstrate that the heuristics of exploratory search significantly help the evaluators achieve an exploratory search system assessment. However, further evaluations would offer interesting feedbacks on the method's usability. For example, the method was not tested in real conditions with a real team of computers scientists evaluating their exploratory search system with the protocol exposed in Section 4.5.2. We need some information about the protocol's understandability to propose more explanation for its improvement.

In Section 3.5.3, we introduce a second version of the model used in this analysis. This new version offers new observed transitions between the model's features. These new transitions do not imply the addition of new heuristics of exploratory search. Indeed, they correspond to heuristics that already exist. For example, transitions H, I $\rightarrow$ B refers to the heuristic *New search launchable from search bar/any element of the interface* and transition D $\rightarrow$ H is facilitated with the heuristic *Bookmarking from result list*. Further analysis of exploratory search sessions on different exploratory search systems may reveal new transitions and exploratory search behaviors implying the design of new heuristics.

4.6 Conclusion

In this section we presented our inspection method of exploratory search systems. This evaluation method consists of a set of heuristics and a procedure for using the heuristics. The heuristics, and the form format we propose, are designed to identify features, interfaces elements or possible actions that support exploratory search behaviors. The evaluation of the heuristics by computer scientists demonstrated that the method significantly help the evaluators identify more and more systematically the presence or the absence of elements or possible actions required to effectively support exploratory search behaviors than without the heuristics. We mentioned in the Chapter 3 the importance of a good usability. Our experiment support this affirmation by showing that a more complete evaluation of these exploratory search systems requires a combined evaluation of the proposed heuristics with well-known

usability and user experience heuristics such as Nielsen and Bastien's and Scapin's ones.

In the evaluation of the proposed method, we noticed that the interrogative form offer an easier experience of the inspection. In order to support the evaluation task we propose a form format of the heuristics and an online evaluation checklist. Users of the method need support in their evaluation tasks of exploratory search systems. Indeed, computers scientists are not familiar with inspection methods. One of the issue with the online form format is that evaluators have to switch between two browser tabs: the evaluation checklist and the inspected system. A split-screen display would allow a parallel presentation of the tabs but it is not always possible. A tool allowing the evaluation without switching between the two different tabs will make the evaluation less difficult. In Chapter 6 we propose such a tool: CheXplore, a Google Chrome Plugin that offers an easier evaluation experience.

In the next chapter we will introduce the second model-based method we design for the evaluation of exploratory search system. This method is an empirical method which propose an adaptable protocol for user tests and an results' evaluation methodology.

Elements of a user testing method 5

5.1 Introduction: the need for an empirical method

In this section, we present the second of our model-based methods for helping exploratory search systems designers (computer scientists) evaluate exploratory search systems. This method complements the inspection method presented in Chapter 4; it allows for a more complete evaluation of a given system.

In contrast to the first one, this second evaluation method is empirical and involves real end-users of exploratory search systems who test a functional prototype or version of an exploratory search system. Just as the first method, this method is based on the model of exploratory search exposed in Chapter 3. It means that the user tests aim to assess if the users indeed perform an exploratory search process. The goal is to find in their exploratory search session the features and transitions of the model of exploratory search we designed. Indeed, in Chapter 3 we observed among other things that the analysis of users performing an exploratory search session on three different exploratory search systems shows relevant information about usability issues and participants' difficulties in their exploratory search task. In other words, users' comments and activity reveal that sometimes users had difficulties to perform an action corresponding to a feature of the model for different reasons such as usability issues and/or missing system's features. These exploratory search sessions showed how real end-users interact with a selected exploratory search system, and the analysis shows relevant information for the improvement of the system's performance in terms of explorability. All these information led us to design an empirical method with a similar procedure as in Section 3.4: we want the participants to achieve their exploratory search task with a certain freedom.

In this section we focus on two elements of a protocol for user testing, in line with what we said previously. We first present a customizable user testing procedure as a context of use of these elements. Second, we introduce the first element: a protocol for the elaboration of exploratory search tasks. Indeed, in an exploratory search session users follow their interest, and their choices and pathways cannot be identical from one user to another. A user can change her goal and her queries multiple times in one exploratory search session, and we cannot anticipate such a behavior because it depends on users' experience and knowledge, state of mind, etc.

We have to provide a certain form of freedom to give the possibility to explore and change goal as they like. An exploratory search task has, consequently, very specific criteria to produce such exploratory search behaviors. The protocol we propose aim to facilitate the elaboration process of such ecological and suitable tasks. Third, we introduce a video analysis grid which helps the evaluators to analyze the exploratory search sessions recorded with the users' comments. This grid aims to facilitate the analysis process of these exploratory search sessions.

In this section, we use the term *task* for designating the tasks proposed to the users at the beginning of each users test.

5.2 A customizable test procedure for an empirical evaluation of exploratory search systems

The two elements we focus on in this section are part of a customizable user testing procedure. We present here the procedure as their context of use. As mentioned previously, the user testing procedure follows the general lines exposed in the evaluation protocol of the model in Chapter 3: a user explores a subject for ten minutes on an exploratory search system. The recorded and commented exploratory search sessions are analyzed with a given video analysis grid based on our model of exploratory search. In Section 2.2 we observed that exploratory search systems are really different in terms of interfaces or databases. Indeed, we cannot propose the same tasks for all the exploratory search systems because they are really different in terms of exploration experience (interfaces and possible interactions) and the explored subjects are limited to the used database. These elements are really important because they make impossible the elaboration of one unique test procedure for all the exploratory search systems. Therefore, we want to propose a *customizable procedure*: a procedure which explains how to design ecological exploratory search tasks and how to analyze the recorded exploratory search sessions.

The customizable procedure can be used on any exploratory search system: the discovery scenario and the exploratory search tasks are system-specific and can be specified by the designers and evaluators of the system. The evaluation procedure can be applied on any exploratory search system because it does not specific to an interface but is more focused on the users' behaviors and exploratory search process.

We describe the customizable user test procedure as follows:

1. **Customize and prepare your test procedure**
   a) **Design one or several exploratory search tasks with the protocol exposed in Section 5.3**
   b) **Prepare a discovery scenario of your system:** an interactive demo presenting the exploratory search system and all its features (e.g. Table 3.7).
   c) **Recruit your participants:** Sova and Nielsen give in [67] "*tips and tricks*" for recruiting users as participants of tests. This book is designed for usability tests but the identification of user(s) profile(s) and their common characteristics is still relevant for this test. These characteristics can be used as filters to select your participants (e.g. same professional activity, same hobbies, age...). Moreover, Nielsen states in [48] that only five users are enough because they can cover 75-80% of the issues in the interface design. We recommend to repeat the users tests procedure with five participants between two improvements of the tested prototype or system.
2. **The test**
   a) **Discovery step of the exploratory search system**
      i. Presentation of the goal of the test
      ii. Discovery scenario : thanks to interactive demo, users learn how to use the system and do not waste time to discover it in the exploratory search session.
   b) Exploratory search session (about 10 minutes)
      i. Presentation of exploratory search task(s): if you have a set of tasks, let each participant choose the most interesting one for her.
      ii. Test session (about ten minutes): The participants performs their exploratory search. Each search session must be recorded with a screen recorder (e.g. Silverback or Apowersoft).
3. **Records analysis:** analyze the results with the evaluation protocol described in Section 5.4.

5.3 Criteria and protocol for the design of exploratory search tasks

In the Chapter 3 we said that users' engagement in a user test is really important, especially for exploratory search systems. To induce this engagement, the users have to be interested in the explored subject. Here we explain how we designed and evaluated our protocol for the elaboration of exploratory search task for users tests.

5.3.1 Task design criteria: a short state-of-the-art

In [66], the authors designed a set of search tasks (including look-up and exploratory search). Exploratory search tasks were designed without taking into account the participants' interest. In their exploratory search task design methodology they follow the *rules* stated in [30] by Kules, Capra and Robert; an exploratory search task:

- Indicates uncertainty and ambiguity in information need
- Suggests knowledge acquisition, comparison, or discovery
- Provides a low level of specificity about: the information necessary and how to find the required information
- Provides enough imaginative context in order for the study participants to be able to relate and apply the situation

Despite these advice and indication, their exploratory search tasks were sometimes too prescriptive. For example, *"Mr Johnson and his family are planning a trip to Paris. Please write a summary of the following: What cultural events are there (for example in August 2010)? What sightseeing to do? Cover also hotels, flights, traveling to Paris, weather*" is an exploratory search task which indicates what kind of information the user has to search: an event, the sightseeing, the hotels, etc. Moreover, the identification to Mr Johnson is not possible even if the user is not a woman. Indeed, for the design of an *ecological task*[1], the task given to the participants has to be realistic. In other words, it should correspond to users' search habits, which they may do on their own in a personal or working context. The task should have involved the users, for example : "*You are planning a trip to Paris. You want to know what cultural events are there (for example in August), and have information on the accommodation or transports*". Moreover, the task can be even less directive "*You will go to Paris in August, you go on [exploratory search system's name] to plan your trip*".

In [4], the users have imposed tasks which were designed without taking into account their personal interest but they align with users' professional activity. However, even if they design tasks in line with participants' professional activity (they are researchers in computer sciences) and in line with Kules and Capra's rules [30], there is still a high probability that they will not do such an exploratory search task on their own, so the ecological aspect is not filled (e.g. "*Collect literature to write a short essay describing similarities and differences between Supervised learning and Unsupervised learning (or Transfer learning and Multitask learning*").

[1] i.e. a task which suits the user's profession and/or interest to be realistic

In [31], the authors give search tasks to their participants, who are undergraduates and graduate students of the University of Maryland at College Park. The exploratory search ones were too prescriptive and their ecological aspect is debatable. For example, the following task is really domain-specific and the participants have to achieve an exploratory search task that they will probably never do on their own: "*Your professor wants you to write a paper comparing the textile industry in three countries in three different continents. Use the catalog to find three countries which have a textile industry about which books have been written. Find three books for each country*".

All these elements show that designing an ecological exploratory search tasks can be really difficult [31]. Kules, Capra and Robert, indicate in [29] a more complete version of what exploratory search tasks should be. Indeed, they spycify it should:

- Indicate uncertainty, ambiguity in information need and/or need for discovery
- Suggest a knowledge acquisition, comparison, or discovery task
- Be an unfamiliar domain for the searcher
- Provide a low level of specificity about:
  - The information necessary for their search
  - How to find the required information
  - How to recognize the required information
- Be a situation which the test persons can relate to and in which they can identify themselves
- Be a situation that the test persons find topically interesting
- Be a situation that provides enough imaginative context in order for the test persons to be able to relate and apply the situation.

5.3.2 Our main task design criteria

These elements were used as a basis for the elaboration of a list of the characteristics that the exploratory search tasks must have to be considered as actual exploratory search tasks, in our context of an exploratory search session evaluation. Indeed, the tasks elaborated with our protocol for the design of exploratory search tasks should check all the following characteristics.

From the previous list we remove the first and the third characteristics. The first one was already included in other characteristics we propose (1.a and 1.b), and the third one because it cannot be used in our protocol for the design of exploratory

search tasks. Indeed, an expert can explore a known domain with an information monitoring approach for example.

The following list of exploratory search tasks' characteristics is used for the elaboration of our protocol and also in its assessment in order to check if the tasks designed by the participants are effectively exploratory search ones.

The exploratory search tasks:

1. Do not provide a clear criteria on:
   a) how to begin the search
   b) when to end the search
2. Suggests a knowledge acquisition, comparison, or discovery task
3. Provides a low level of specificity about:
   a) the information necessary for their search
   b) how to find the required information
   c) how to recognize the required information
4. Must be ecological:
   a) a situation which the test persons can relate to and in which they can identify themselves
   b) a situation that the test persons find topically interesting
   c) a situation that provides enough imaginative context in order for the test persons to be able to relate and apply the situation

5.3.3 Additional task design criteria based on Bloom's taxonomy

Our protocol is based on a set of exploratory search tasks characteristics exposed in the last section, but also on Bloom's taxonomy. Indeed, Marchionini exposed in [33] a set of exploratory search activities based on Bloom's taxonomy (see Figure 3.5), such as knowledge acquisition, analysis... The activities induced by the tasks must be in line with the Bloom-based activities.

Bloom's taxonomy of educational objectives [10] was created in 1956 and promotes higher forms of thinking in education. It is mainly used by teachers to ask the right questions to identify the level of understanding of their pupils or students. In order to help them, some kits or tables are available on the Web or on books. These kits and/or tables present the main Bloom's objectives, the same mentioned

by Marchionini in his model, such as knowledge acquisition, analysis, synthesis, interpretation of ideas or comparison or aggregation of data and concepts, etc.

Several kits/tables link Bloom's educational objectives to keywords, examples of tasks, actions or questions. We use a kit available on the Web[2] to the design of our protocol. Indeed, we use the proposed keywords (e.g. *ask*) or actions (e.g. *classifying*) and the possible outcomes (e.g. a *list*) in order to guide and facilitate the tasks design process.

5.3.4 Protocol for the design of exploratory search task

As a reminder, we design this elaborating exploratory search task protocol for computer scientists who are both the designers and the evaluators of a given exploratory search system. We want a step by step protocol, as simple as possible. With this goal in mind, the keywords and the activities exposed in these kits of Bloom's taxonomy are really useful because they can be used for the design of the search goal. From the kit of Bloom's taxonomy, we use the keywords, the possible outcomes and the actions.

From the kit, we list these relevant terms which can be used in our protocol. For example, in the kit the different outcomes inspired us to propose our own list of outcomes that could be used by the evaluators in their task design: (make a) list, presentation, demonstration, plan, survey, examples, summary... In step 2.2 we ask them to select one outcome that the user will do in her search session.

In the following Table 5.1 we present our protocol of designing exploratory search tasks. We add in parallel an example of the outcome obtained for the exploratory search system Discovery Hub. The protocol is divided into two parts: first, the elaboration of the task (topic, goal and context design) and secondly the verification that the elaborated task is ecological and respects the characteristics we proposed in Section 5.2.

We use the keywords, actions and outcomes of the Bloom's taxonomy kit in the step 2.2 which consists of defining for what purpose the user will do an exploratory search task: *discover* the topic, find results *similar* to the topic, make a *list / plan / presentation /* paper /report, prepare of *organize* an event or a trip, give *example*, *compare* two different concepts, *find* information... This list can be used or adapted by the evaluators as they like.

[2] e.g. in `https://goo.gl/WuxacF`

Protocol's steps	**Outcome example for Discovery Hub**
STEP 1: Define a topic to explore by the users on your exploratory search system.The used database and/or the goal of your system can guide you in this topic definition.	Discovery Hub is based on DBpedia (a open knowledge graph of Wikipedia developed by the Semantic Web community: `https://wiki.dbpedia.org/about/`). With this system, a user can explore a wide range of subjects (countries, music, politics...). In this example we choose the topic *Senegal*.
STEP 2: In this step, you will define the goal of the exploratory search task **STEP 2.1:** Indicate why the user will explore the selected topic. Choose one of the these activities : • Elaborate something (e.g. a text, a presentation, a plan...). • Learn more about the subject, linked to the users' personal or professional interests. • Share information on the subject in order to: recommend, advice, teach, explain, etc. **STEP 2.2:** Indicate what the user will do to achieve her goal. For example: discover the topic, find results similar to the topic, make a list / plan / presentation / paper / report, prepare of organize an event or a trip, make a diagram, give example, compare two different concepts, find information on the selected topic (e.g. bookmark results), compare...	*STEP 2.1:* Learn more about the Senegal *STEP 2.2:* make a list

Continued on next page

Tab. 5.1 – *Continued from previous page*

Protocol's steps	**Example with Discovery hub**
STEP 2.3: From the previous elements, write down the task and give a context of search. Ask yourself why the user will do such an exploration. This context will allow an easier identification and appropriation of the exploratory search task. It has to be plausible and realistic with regards to the user's life and habits.	*STEP 2.3:* "A friend from Senegal invited you to visit him in Dakar. To prepare your trip, you want to list the interesting things to see and discover"
STEP 3: Make sure that the task is *ecological*. In other words, the task should represent a real-life situation or a simulation of daily activities. The explored topic should suit the user's interest and/or profession (e.g. "write an article on a topic" should be right only for journalists or bloggers because it is in line with their daily activity and profession).	*STEP 3:* It is OK. Our task written in step 2.3 presents an invitation from a friend. It is a common event which can concern anybody.
STEP 4: Make sure that the task won't contain clear criteria on when to end the search. For example, do not provide limits for the duration of the search session and the number of results the user have to find ("find three Italian restaurants" vs. "find Italian restaurants"). You only have to give a task and explain why the user have to do it.	*STEP 4:* The task is ecological. There is no time limit or specified number of results requested for the list.
STEP 5: Make sure that the specification on how to find the information or how to evaluate them is kept to the minimum, or not specified at all. In this sense, the user should not have a clear idea on how to begin the search.	*STEP 5:* This is indeed the case. We did not specified what to search (places, food, museums, event...). The user may interpret and appropriate the task.

Tab. 5.1: Protocol for the design of exploratory search tasks

5.3.5 The evaluation of the exploratory search task design protocol

In the evaluation of the protocol exposed in Table 5.1, we want to verify if users can design exploratory search tasks with the given indications. We want to evaluate the following hypothesis:
H1: with our protocol for the design of exploratory search tasks, the users design effectively exploratory search tasks.

5.3.5.1 Method

We selected twenty participants, all of them are computer scientist and did not know anything about exploratory search before the test. Ten participants have to design exploratory search tasks for Discovery Hub, and the ten others for 3cixty. Indeed, the elaboration of exploratory search tasks takes place in the context of an evaluation of a specific exploratory search system, and more precisely in the evaluation phase of a functional prototype or system. We wanted to test our protocol with different exploratory search systems.

The evaluation follows this procedure:

1. Presentation of the test's goal
2. Presentation of the system (Discovery Hub or 3cixty) with discovery scenarios exposed in Table 3.7 in Chapter 3.
3. Elaboration of two exploratory search tasks for one of the two systems by using the protocol exposed in Table 5.1
4. End of the test and acknowledgment

5.3.5.2 Results

As mentioned previously, the designer of the protocol evaluates the participants' tasks with the list of exploratory search tasks' characteristics exposed at the end of Section 5.2. Indeed for each task we verify if it checks all the characteristics.

In Table 5.2 we give an example of the analysis with one of the participants' tasks exposed in Table 5.3. For all the participants, 100% of the tasks which where designed with our protocol are effectively exploratory search tasks. We also noticed that the tasks were designed without notable difficulties.

Then the hypothesis H1 is validated: our protocol for the design of exploratory search task is validated.

5.3.5.3 Conclusion and perspectives

The evaluation shows that the protocol we designed allows the elaboration of exploratory search tasks. In this protocol we propose to design a non customizable task. Depending on the system's database, the users of the method (i.e. the designers of the systems or evaluators) can also design with our protocol for the design of exploratory search tasks *templates of tasks* which can be customized in the last minute with the participants. This solution can imply a more important engagement from the participant because it may suit its personal interest, and excite her curiosity.

For example, Discovery Hub offers a large range of possible subject to explore. The task "*you heard on the radio a song of XX, you go on Discovery Hub in order to know more about her/him*" can suit the participant's musical taste. The same task with an imposed artist would not be as attractive to the participant. Further evaluations are necessary to find out whether our protocol allows the design of such customizable template tasks.

5.4 A protocol for the model-based video analysis

In Section 5.2, we present a procedure of customizable users tests. We present in Section 5.3 a way to design suitable and ecological exploratory search task with a given protocol.

In this section we present a video analysis grid for the evaluation of recorded exploratory search sessions. It corresponds to step 3 of the procedure presented in Section 5.2: *the records analysis*. In this evaluation step, the evaluators analyze records of exploratory search session: screencasts and users' comments (they explain what they do and why). As the heuristics of exploratory search, our video analysis grid is based on our model of exploratory search process presented in Section 3. Indeed, the evaluators identify in the video records and users' comments which model's feature the user is doing. The video analysis grid we propose aims to facilitate this identification.

With such an analysis, the evaluators see if the users achieve all the features without difficulties, and if it is not the case, they can identify the source of the issues (e.g. missing features, usability issues...).

Our task design criteria	**Our evaluation of the task:** *"you planned a trip in Corsica with a limited budget, you go on 3cixty in order to plan and organize this trip"*
The task does not provide a clear criteria on: • how to begin the search • when to end the search: she does not specify limits in terms of time or number of results.	This is indeed the case: the task is not limited to the hostels or places of interest for example. Furthermore, it does not indicate any limit in term of time or number of results.
The task suggests a knowledge acquisition, comparison, or discovery task	This is indeed the case: the task suggests learning, comparing and discovering activities. Indeed, planning a trip may imply to discover the places of interest, compare different hostels,etc.
The task provides a low level of specificity about: • the information necessary for their search • how to find the required information • how to recognize the required information	The task provides effectively a low level of specificity about required information. The only indication which will be used as a framework of the search is the limited budget. Apart from that, the task does not specify what kind of information/results the user has to search.
The task must be ecological: • a situation which the test persons can relate to and in which they can identify themselves • a situation that the test persons find topically interesting • a situation that provides enough imaginative context in order for the test persons to be able to relate and apply the situation	The task gives a situation providing effectively an imaginative context. The user can assume ownership of the task and the way they want to explore the topic. The situation of planning a trip is a common situation and generally interesting for every body. For all these reasons, we can say that the task is ecological.

Tab. 5.2: The evaluation with our task design criteria of the task proposed by Participant 1 on 3cixty.

Discovery Hub	**3cixty**
Participant 1: *You organize a birthday party and you go on Discovery Hub to prepare a playlist*	Participant 1: *You planned a trip in Corsica with a limited budget, you go on 3cixty in order to plan and organize this trip*
Participant 2: *You want to establish a list of movies in order to prepare a movie night with friends and you go on Discovery Hub*	Participant 8: *You are going to make a business travel in Paris. You want to occupy your freetime between two meetings and you go on 3cixty to explore the museums' exhibitions in that period*
Participant 6: *You are interested in the field of Psychology and you go on Discovery Hub in order to compare, discover the differences and similitudes between the subdomains (cognitive psychology, clinical psychology...)*	Participant 11: *You planned a trip in Barcelona and you want to know more about the city (customs, traditions...)*

Tab. 5.3: Examples of tasks designed by using our protocol for the elaboration of exploratory search tasks

The element of method we propose here is a video analysis grid that helps the user in the identification of the model's features in a user's exploratory search session and comments. Such an analysis highlights the difficulties encountered by the participants and gives clues of improvements.

We first explain how the grid was designed and, in a second time, present the first results of the evaluation of the protocol.

5.4.1 Design method of the video analysis grid

As mentioned previously the objective of the grid we want to design is to facilitate the evaluators in their video records analysis of exploratory search sessions associated with users' comments. Indeed, as a reminder, the evaluators will be computers scientists and they are not familiar with users tests evaluation, they need an easy approach of the video analysis activity.

Supporting their evaluation task implies a video analysis grid that eases the identification of the features in the records of exploratory search sessions. The identification of the features means, the ones who the user actually uses in the exploratory search session, and also the failed features. The identification of the source of the failure

allow the evaluators to improve the system's interface of interactions in order to address them.

Table 5.4 shows the different design phases of the adopted approach for the wanted video analysis grid. Indeed, it is a description of the model's features in terms of action/comments and questions. In the design phase of the grid, we first detail for each features the users' actions and/or comments they can make. From these actions, we design secondly a set of corresponding questions for each feature.

The questions related to one model's feature make the evaluation easier. Indeed, when an evaluator is analyzing an exploratory search session of a participant of the test, and can answer "*yes*" to one of the questions, it means that the participant is doing the associated feature.

An example with an real participant's comment[3] "*I zoom on the map [on Frankenplace] and briefly seek if there is no other castles that I could know*". from this user's comment and activity, the evaluator can only answer *yes* to the questions corresponding to step H (Browse results) of the model:

- Does the user scan the result list?
- Does she want to know what kind of results are given by the system?

Indeed, the participant's comment shows that she is actually scanning the results list ("*briefly seek*") and explore and discover the different retrieved results with a vague goal in mind.

5.4.2 Evaluation of the video analysis protocol

5.4.2.1 Evaluation goal and hypotheses

The goal of this evaluation was to test the grid's capacity to help the participants in the identification. We designed first a set of short video clips of a fictive exploratory search session and fictive associated user's comments. The scenario of these video are inspired from exploratory search sessions analyzed in Section 3.4 and 3.5. As in the procedure exposed in Section 5.2, step 3, the task of the participant is to identify which feature the fictive user is actually using in the video clips.

For a first evaluation of the grid, we choose to propose videos with easily identifiable model's features. Indeed, we do know that our participants may, at best, have a

[3] From a recorded exploratory search session on Frankenplace, in the context of the model's evaluation in Section 3.5.

low level of knowledge of exploratory search concept. Furthermore, we actually know that the video analysis here is quite artificial, because it does not represent a real video analysis of an exploratory search session. However, with a first test of the video analysis grid as we propose here, even with non-experts of exploratory search concept, we have interesting results of the understandability of the grid and the questions.

Then the objective of the test is to analyze the capacity of the video analysis grid to help the participants in the identification of the features in the given video clips. We do know that further evaluation are already needed. From the beginning, we do not expect a perfect identification from the participants because we know that a non-expert of exploratory search concept may have more difficulties than experts. In the case of the participant identified incorrectly the features, we seek reasons of the failures and possible solutions.

5.4.2.2 Evaluation method

Participants.
For this evaluation of our video analysis grid, we selected twenty participants. All of them are computer scientists: twelve have a low level of knowledge on exploratory search concept because they participated to the other users tests exposed in Sections 4.5 and 5.3 (participants 1, 2, 3, 4, 5, 6, 11, 12, 13, 15, 17, and 20).The eight other did not know anything about exploratory search before the test (participants 7, 8, 9, 10, 14, 16, 18 and 19).

The participants were asked to perform online (using a Google Form) an analysis of the videos of users carrying out an exploratory search task on two different exploratory search systems:

- First group of ten participants (Group 1) was asked to analyze videos on 3cixty: `https://goo.gl/cPZFMN`
- Second group of ten participants (Group 2) was asked to analyze videos on Discovery Hub: `https://goo.gl/4DF3Fz`

Tab. 5.4: Video analysis grid

Model's feature	Users' actions and/or comments	Questions
A	The user specifies her goal(s) of search, expresses her lack of knowledge, her information needs. The user specifies how she decided to start the search	• Does the user's express her goal(s) of search? • Does she express her information need or lack of knowledge? • Does the user specify how she decides to start the search? • Does she express her angle of attack of the problem?
B	The user uses the search bar or the filters tools in order to formulate or reformulate her query	Does the user formulate the query with filter tools or search bar or any other means?
C	The user says that an information (on the result list or on an element description) is interesting. Never mind if this information is relevant or not for the current problem and search.	Does the user say or imply that some information is interesting (never mind if this information is linked or not to the current search or problem)?

Continued on next page

Tab. 5.4 – *Continued from previous page*

Model's feature	**Users' actions and/or comments**	**Questions**
D	The user adds to bookmarks one or several results with some memorization features, from the list of result or one element description.	Does the user add to bookmarks (or other memorization tools) or write down on a paper relevant information?
E	The user wants to know more about an element (query, result) or the link between the query and a result. She may use different descriptive information or features, e.g. a descriptive text, pictures, tags, map, links, etc.	Does the user examine one element (result or query) in order to: • discover it? • check an information about it? • or to know more about the link between the query and the result?

Continued on next page

Tab. 5.4 – *Continued from previous page*

Model's feature	Users' actions and/or comments	Questions
F	The user says that she wants to change the goal of search and maybe explain why : it can be a call into question of the current goal and query, or maybe because she finds something more interesting in her exploratory search process (e.g. a relevant result, or surprising one...)	• Does the user say that she does not find the current goal or the query relevant anymore? • Does she says that she finds something more interesting for her exploration (e.g. a relevant result or surprising one)? • Does she simply say that she wants to change the goal or the angle of attack?
G	The user can use the browsing history, the breadcrumb or the back/next buttons.	• Does the user uses the browsing history, the bread crumb or the back or next buttons? • Does the user step backward or forward? • Does the user go back to a previous screen?

Continued on next page

Tab. 5.4 – *Continued from previous page*

Model's feature	Users' actions and/or comments	Questions
H	The user browses or scans the results given by the system.	• Does the user scan the result list? • Does she want to know what kind of results are given by the system?
I	The user analyses a result description. She evaluates its relevance based upon the goal of search.	Does the user evaluate the results' relevance for her goal of search (e.g. the user expresses that a result is relevant or not and why) ?
J	The user stops her search session.	Does the user stop her search session?

Tab. 5.4: Video analysis grid

Continued on next page

Materials.

Videos and scenarios. We designed first a set of short video clips of a fictive exploratory search session and fictive associated user's comments. The scenario of these video clips are inspired from exploratory search sessions and users' comments analyzed in Sections 3.4 and 3.5. The video clips and the associated user's comments match with actual possible exploratory search behaviors.

As mentioned in the last section, we did not propose features difficult to identify. Each video can illustrate one to three features. Indeed, even if the transitions between the features are not central in the evaluation procedure, they are part of the exploratory search process and behaviors. We want to propose realistic video clips of exploratory search sessions which correspond to a certain extent to the real analysis of ten minutes of exploratory search sessions. Thus, we have to design video clips which illustrate sometimes one or several transitions.

Table 5.5 presents the different fictive user's comment which are used in the elaboration of the video clips (screencasts). We used Apowersoft[4], a free online recorder, for the elaboration of the videos clips.

Video analysis grid/form. To facilitate the analysis of the videos, we propose a printed version of the Table 5.4 and procedure, available on Figure 5.1. The participant can analyze the video and read the grid and the questions in parallel of the form.

For each video we ask the participant:

1. specify which feature(s) the fictive user is performing on the video clip
2. specify on a Likert scale if the choice is based on the video (user's actvitiy)
3. specify on a Likert scale if the choice is based on the user's comment

5.4.2.3 Procedure

In the recruiting phase of the participants, we present briefly the exploratory search system they will analyze. We only indicate what kind of exploratory search tasks the users can achieve on such system :

- By using 3xicty exploratory search system, the user can organize trips or outings by exploring the different places (museums, restaurants, etc.) and events of a city (here, Nice).

[4] `https://www.apowersoft.com/free-online-screen-recorder`

What is the user's activity?		Name of the step process	Do you understand the questions?		
			Yes	Partially	No
• Does the user's express her goal(s) of search? • Does she express her information need or lack of knowledge? • Does the user specify how she decides to start the search? • Does she express her angle of attack of the problem?	A	**Define the search space**	☐ ☐ ☐ ☐	☐ ☐ ☐ ☐	☐ ☐ ☐ ☐
Does the user formulate the query with filter tools or search bar or any other means?	B	**Query (re)formulation**	☐	☐	☐
Does the user say or imply that some information is interesting (never mind if this information is linked or not to the current search or problem)?	C	**Information gathering**	☐	☐	☐
Does the user add to bookmarks (or other memorization tools) or write down on a paper relevant information?	D	**Put some information aside**	☐	☐	☐
Does the user examine one element (result or query) in order to: • discover it? • check an information about it? • or to know more about the link between the query and the result?	E	**Pinpoint search**	☐ ☐ ☐	☐ ☐ ☐	☐ ☐ ☐
• Does the user say that she does not find the current goal or the query relevant anymore? • Does she says that she find something more interesting for her exploration (e.g. a relevant result or surprising one)? • Does she simply say that she wants to change the goal or the angle of attack?	F	**Change of goal(s)**	☐ ☐ ☐	☐ ☐ ☐	☐ ☐ ☐
• Does the user uses the browsing history, the bread crumb or the back or next buttons? • Does the user step backward or forward? • Does the user go back to a previous screen?	G	**Backward/forward steps**	☐ ☐ ☐	☐ ☐ ☐	☐ ☐ ☐
• Does the user scan the result list? • Does she want to know what kind of results are given by the system?	H	**Browsing results**	☐ ☐	☐ ☐	☐ ☐
Does the user evaluate the results' relevance for her goal of search (e.g. the user expresses that a result is relevant or not and why) ?	I	**Results analysis**	☐	☐	☐
Does the user stop her search session?	J	**Stop the search session**	☐	☐	☐

Fig. 5.1: The video analysis grid presenting the model's features associated to corresponding questions. We ask the user to specify their understandability of the questions.

- By using Discovery Hub exploratory search system, the user can explore and learn new things on a wide range of subjects (countries, music, politics...).

For the persons who have agreed to participate in the test, we give a printed version of the video analysis grid (see Figure 5.1). After that, we send an e-mail which explains the procedure and includes the URL link to the online form. We ask them to:

1. Read carefully the given table (Figure 5.1) describing the model with the associated questions.
2. Mention if they understand the questions totally (yes), partially or not at all (no). We also ask them to give us back the paper after the test.
3. Click a URL link to access the online test
4. Follow the form

In the online form, we propose several short video clips of a search session on an exploratory search system, 3cixty or Discovery Hub. We ask the participants, for each video, to find in the user's actions and comments the corresponding model's feature. If the participant does not find a feature, she can select two different options:

- Other activity (precise why in comment)
- Any of them (not found)

We observe in the results analysis that none of the previous options were selected by the twenty participants.

Tab. 5.5: List of user's comments on Discovery Hub and 3cixty linked to our model's features of exploratory search. These comments are used in the elaboration of screencasts.

	Feature	Discovery Hub	3xicty
Video 1	A	"Later, I heard about Bob Dylan on the radio, and I go on Discovery Hub in order to discover him"	"Some friends come to visit me in Nice, I go on 3cixty to organize their trip..."
	B	"I write on the search bar « Bob Dylan » and launch the search."	"... I want to see museums that I can suggest them. I clic on « Find an interesting place » and select « Arts & Entertainment »"
Video 2	H	"I am scanning the result list in order to discover what kind of results the system proposes"	"I am scanning the result list in order to discover what kind of results the system proposes"
Video 3	E	"I click on « Lay Lady Lay » to discover the link between the query [Bob Dylan] and this result"	"I am reading the description of « Musée Masséna »..."
	G	NA	"... and I use the bread crumb to go a step backward"
Video 4	C	"This result seems interesting..."	"This restaurant seems really good..."
	I	"... but not really relevant for my goal of search..."	"... but too far from the museum."

Continued on next page

Tab. 5.5 – *Continued from previous page*

	Feature	**Discovery Hub**	**3xicty**
	G	"... I close the pop-up to go back the last screen, the results' list "	NA
Video 5	F	"When I saw the result « John Lennon » I thought launching a new search on the Beatles is more interesting for my goal of search..."	"At this instant I thought that it would be more interesting to organize the trip focusing on « Place Rossetti »... "
	B	"... then, I write on the search bar « The Beatles »"	"...that is why I open a new box and use the pivot button in order to discover nearby restaurants"
Video 6:	D	"I bookmark this result for other search session"	"I bookmark this result for further search session"
Video 7	J	"Here, I close the window because I do not have the time to continue."	"And then, I stop the search, I will continue it tonight after work"

Tab. 5.5: List of user's comments on Discovery Hub and 3cixty linked to our model's features of exploratory search. These comments will be used for the elaboration of screencasts.

Continued on next page

5.4.2.4 Results and discussion

For each participant, we analyze the features correctly assigned and those which are not. The results are exposed in Table 5.8. We sum up them in an histogram on Figure 5.2 focusing on Group 1, Group 2 and all the participants' score.

We observe that the Group 1 on Discovery Hub has a higher score of correct assignation than Group 2 on 3cixty, respectively: 87.3% and 70%. We also observe that all participants correctly assigned at 78.6% the features (overall score). Moreover, with a standard deviation of 16.8 for the all participants (with no distinction between the systems), 9.8 for Group 1 Hub and 18.2 for Group 2, we also note large variations from one participant to another. Indeed, only three participants (1, 9 and 20) made a perfect assignation of the feature, and the poorest score are not linked to the level of knowledge on exploratory search concept. Indeed, participants number 13, 17 have a low level of knowledge but they have a bad score. Moreover, number 13 have the lowest score of correct assignation of the features: 45.5%. Then, we can only say that the level of expertise seems to be an important part of the correct assignation of the features but further evaluations are needed with real experts in exploratory search concept. With these elements we can only hypothesize the reasons of these score and variation between the system or more broadly the participants of the test which can be linked with:

- the level of knowledge on the exploratory search concept and/or the evaluated system
- the less accessible interface and interactions proposed by the two systems that do not imply the same effort of understanding users' activity on the video clips
- the grid which is maybe be difficult to understand.

We cannot explain why the participant assign a feature to the wrong place in the form. Indeed, even the specified understandability of the questions they checked at the beginning of the test (see Figure 5.1) cannot allow an affirmation of a bad formulation of the questions. Indeed, they mostly indicate that they understand the questions, and when they do not, it does not explain all their mistakes.

For example, participant 13 indicates that she does not understand all the questions corresponding to features E, F and I. And she does actually assign Feature E to the wrong place (features B and I), but also Feature C instead of D, I instead of E, A instead of B and vice versa. We can observe on Table 5.7 that features E and I are often confused. The same observation can be done with features A and B. These four features have maybe close definition and questions, and a person with a very

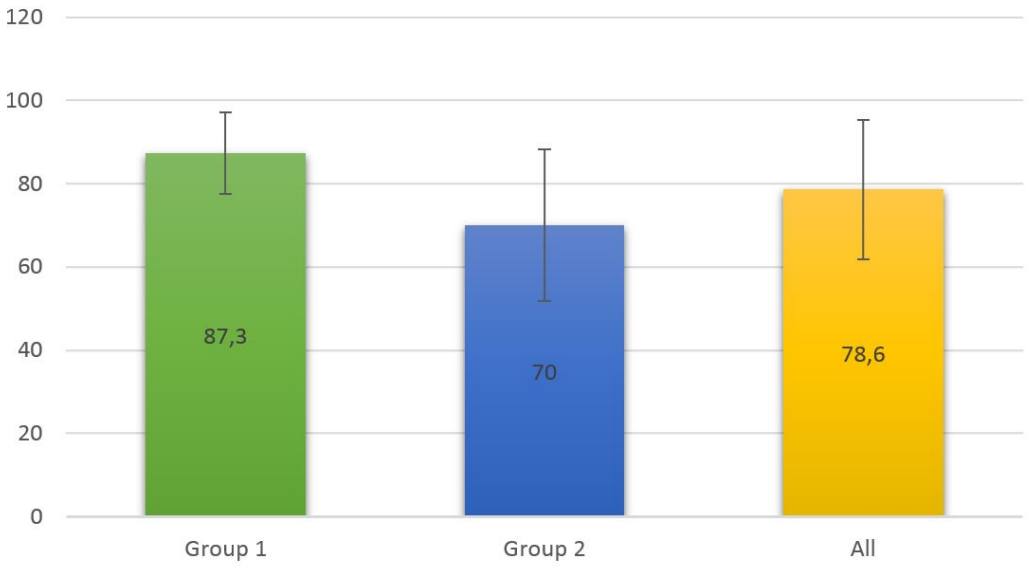


Fig. 5.2: Participants' score in percentage in the evaluation of the video analysis grid for the Group 1 on Discovery Hub, Group 2 on 3cixty, and for all the participants.

low level of exploratory search activity and specificity cannot distinguish easily the differences between the activity they describe, even with our video analysis grid.

For each features, we analyze those correctly assigned and those which are not. The results are exposed in Table 5.6. We can observe that 77.3% the features are correctly assigned. However, we can see that features A, C, E, and I have a low level of correct assignation, respectively: 68%, 64.7%, 54.2%, 56.3%. Table 5.7 shows that some features are confused more than once and consequently incorrectly assigned:

1. Features A and B
2. Feature C, E and I
3. Feature E and H
4. Features F and I
5. Features J and I

If we have no hypotheses explaining cases 3 and 5, we can suppose that cases 1, 2 and 4 reflects concepts which are too close for an easy distinction. Indeed case 1 refers to definition of the goal of search (features A) and the query (features B) which are difficult to distinguish for the person with any knowledge in information seeking.

Case 2 refers to the user's activity on an element of information (simple information or a result): explore (feature E), analyze (feature I) and gather information (fea-

	Expected total	Correctly assigned	Incorrectly assigned	Observed total	% Correctly assigned	% Incorrectly assigned
A	20	17	8	25	68	32
B	40	31	2	33	93,9	6,1
C	20	11	6	17	64,7	35,3
D	20	19	2	21	90,5	9,5
E	20	13	11	24	54,2	45,8
F	20	17	6	23	73,9	26,1
G	20	18	0	18	100	0
H	20	16	3	19	84,2	15,8
I	20	9	7	16	56,3	43,7
J	20	20	3	23	87	13
Average		17,1	4,8	21,9	**77,3**	**22,7**

Tab. 5.6: Results for each model's features.

ture C). Here again, the low level of knowledge in information seeking may imply difficulties in the differentiation of the features.

In case 4, feature F refers to a user changing her goals. In the user's comments, for the two systems, the user says that launch a new search on a new subject would be *more interesting*. We suppose that the participant focused on the term *interesting*, which can be related to features C or I, and not realized that the user was actually changing her goals of search.

5.4.2.5 Conclusion of the preliminary evaluation

In the evaluation of the video analysis grid, we want to evaluate its capacity to help the participants in the identification of the features in video clips on an online form. Despite scores over 70% of good assignation, the results show a high disparity in the participant's score (see Figure 5.2). We cannot affirm the reasons of these results, but we can propose possible explanation for the difficulties the participants faced:

Assigned in place of →	A	B	C	D	E	F	G	H	I	J
A		7				1				
B	1				1					
C				1	2			1	2	
D			1			1				
E	1	1	4					3	2	
F		1							5	
G										
H			1		1		1			
I			3		4					
J							1		2	
Total	2	9	9	1	8	2	2	4	11	0

Tab. 5.7: Details of the wrong assignation of the features.

- the level of knowledge on exploratory search concept was too low for the participants;
- the level of knowledge on the evaluated system: the interfaces are not always easily understandable, and the participant did not understand what the fictive user was doing in the video clips, especially on 3cixty;
- some participants perform the evaluation in a disturbing environment (noise around, late at night...);
- some participants perform the evaluation in haste;
- the online form was not easy to understand and to perform the evaluation;
- the participants do not perceived the difference between the questions associated to the model's features;
- some questions are insufficiently specific of one features, they are sometimes confused.

Further evaluation of the video analysis grid are needed for a complete evaluation. The next evaluation must be more realistic: the participants should have a higher level of knowledge of the exploratory search concept and the evaluated system. The video clips should be longer and real clip from a real exploratory search session, also illustrating some features' failure cases. Moreover, the groups of model's features A

Participant number	Exploratory search system	Correctly assigned	Incorrectly assigned	% Correctly assigned	% Incorrectly assigned
1	Discovery Hub	11	0	100	0
2		8	3	72,7	27,3
3		10	1	90,9	9,1
4		10	1	90,9	9,1
5		9	2	81,8	18,2
6		10	1	90,9	9,1
7		10	1	90,9	9,1
8		9	2	81,8	18,2
9		11	0	100	0
10		8	3	72,7	27,3
11	3cixty	9	2	81,8	18,2
12		10	1	90,9	9,1
13		5	6	45,5	54,5
14		6	5	54,5	45,5
15		9	2	81,8	18,2
16		6	5	54,5	45,5
17		7	4	63,6	36,4
18		8	3	72,7	27,3
19		6	5	54,5	45,5
20		11	0	100	0
	Average	8,65	2,35	**78,6**	**21,4**

Tab. 5.8: Details of the wrong assignation of the features for each participants.

& B and C & E & I must have more distinct questions avoiding any confusion with the other feature(s).

5.5 Conclusion

In this section we focused on two elements of a user testing method: the protocol for the elaboration of exploratory search tasks (Section 5.3) and a video analysis grid for the evaluation of recorded exploratory search sessions associated to users' comments (Section 5.4). These two elements are part of a customizable user test procedure of exploratory search systems proposed in Section 5.2.

The first element, the protocol for the elaboration of exploratory search tasks, is based on our criteria of exploratory search tasks and Bloom's taxonomy. They allow a step by step elaboration of ecological and suitable tasks for any exploratory search systems. The evaluation shows that 100% of the tasks designed by twenty computer scientists are effectively exploratory search tasks. We note that our protocol is easily accessible because none of the participant have knowledge on exploratory search concept.

The second element, the video analysis grid, is based on our model of exploratory search and aims to facilitate the evaluators in their video analysis of exploratory search sessions. The goal of the first evaluation of the grid was to evaluate its capacity to help the participants in the identification of the features in video clips on an online form. The results show a large disparity in the scores of the participants even if the averages are higher than 70%. We underscored the possible reasons and factors that can explain such results, in particular the context of the search (e.g. noise around the participant...), the selected participants and some set of model's features which have questions and definitions that are sometimes confused: features A&B and C&E&I. Further evaluations of the video analysis grid, closer to reality, are needed for a more complete assessment. In other words the next evaluation must involve participants that fit the description of the end-users of the methods (computers scientists who designed an evaluated exploratory search system), and a video of a real user performing an exploratory search task on the system. In addition, the video has to illustrate failed features. Furthermore, the previously mentioned confused set of features must have more distinct questions.

The video analysis need a tool that helps the user in their annotation task. This tool could facilitate the video analysis. With such a tool, the evaluators could annotate a video of an exploratory search session by specifying which model's features the user is actually doing and how it is realized (e.g. succeed, performed with difficulties and

failed). We can imagine that the evaluator could also indicate on her annotation the reasons of the difficulties or failures if it is the case (e.g. usability issue or missing system's feature) export in a Word or PDF document all these elements in order to help the designer to improve the evaluated exploratory search system between two evaluations.

6 Designing a tool for exploratory search interfaces inspection

6.1 Introduction: the need to support the evaluation task

In Chapter 4 we propose an inspection method of exploratory search systems with the design of our heuristics of exploratory search. To ease their use, we also propose an online *evaluation checklist*[1] which is a Google Form using our heuristics in their interrogative form. For each question, the evaluators can record online their answers. Computers scientists are not familiar with inspection methods and this way of presentation makes the evaluation more practical and consequently easier to perform.

However, one of the issues with the online form format is that evaluators have to switch between two browser tabs: the evaluation checklist and the inspected system which does not facilitate their task as much as it could be. A split-screen display would allow a parallel presentation of the tabs but it is not always possible (*e.g.* on a laptop screen). In addition, in heuristics evaluations, evaluators usually take screenshots of the interface illustrative of the identified issues. Our evaluation checklist does not give the possibility to take screenshots and the split-screen display can be a problem.

A tool facilitating the evaluation gives on the same screen/tab the possibility to:

1. inspect the interface,
2. answers to the heuristics of exploratory search [55],
3. take screenshots.

As the evaluation checklist, the possibility of exporting the answers on a report would facilitate the evaluation process.

[1] `http://bit.ly/evaluation-checklist`

In the present chapter we introduce **CheXplore**, a Google Chrome plugin that meets our needs for facilitating the interface inspection of exploratory search systems.

6.2 Which inspection tool

In the user experience and usability domain, several tools already exist for the support of inspection methods, especially heuristics evaluation. The two quite similar Google Chrome plugins we identified are UX Check[2] and Capian[3].

UX Check is a free tool that helps the evaluators identify and annotate on interface's screenshots usability issues through a heuristic evaluation. It also allows the exportation of the evaluation on a Word document. The plugin presents a lateral bar on the left with all the information of the plugin and its settings. UX Check proposes to evaluate the interface with Nielsen's heuristics [47], but also allows the addition of customized heuristics or criteria. For each screenshot, the evaluators can:

1. indicate the involved heuristic
2. add notes
3. add recommendations
4. indicate the issue's severity

The main issue of this tool is that the screenshot is based on the HTML tags of the website. It means that the evaluator is not free in the screenshot to highlight several elements (linked to several different HTML tags), e.g. Figure 6.1. It is very problematic in our context of exploratory search systems because the way they presents the results or data makes UX Check some time unusable (e.g. Linked Jazz[4] is a graph on which the plugin does not work).

Capian's Google Chrome plugin is, unlike the UX Check, a commercial tool. Obviously, such a product implies a more successful user experience: first, there is no lateral bar reducing the website's interface on the screen. It first presents the user-centered design approach an display, and the evaluators are not limited by the HTML tags for their screenshots. Second, it allows a collaborative evaluation and makes easier the team discussion and collaboration. With Capian, the evaluators can inspect the interface with Nielsen's heuristics [47], Bastien and Scapin's ergonomics criteria [5] or add their own criteria. Videos presenting Capian are available on Youtube: `https://goo.gl/t69KGL`. In Capian's plugin, the screenshots present the

[2] `https://www.uxcheck.co/`
[3] `https://capian.co/`
[4] `https://linkedjazz.org/network/`

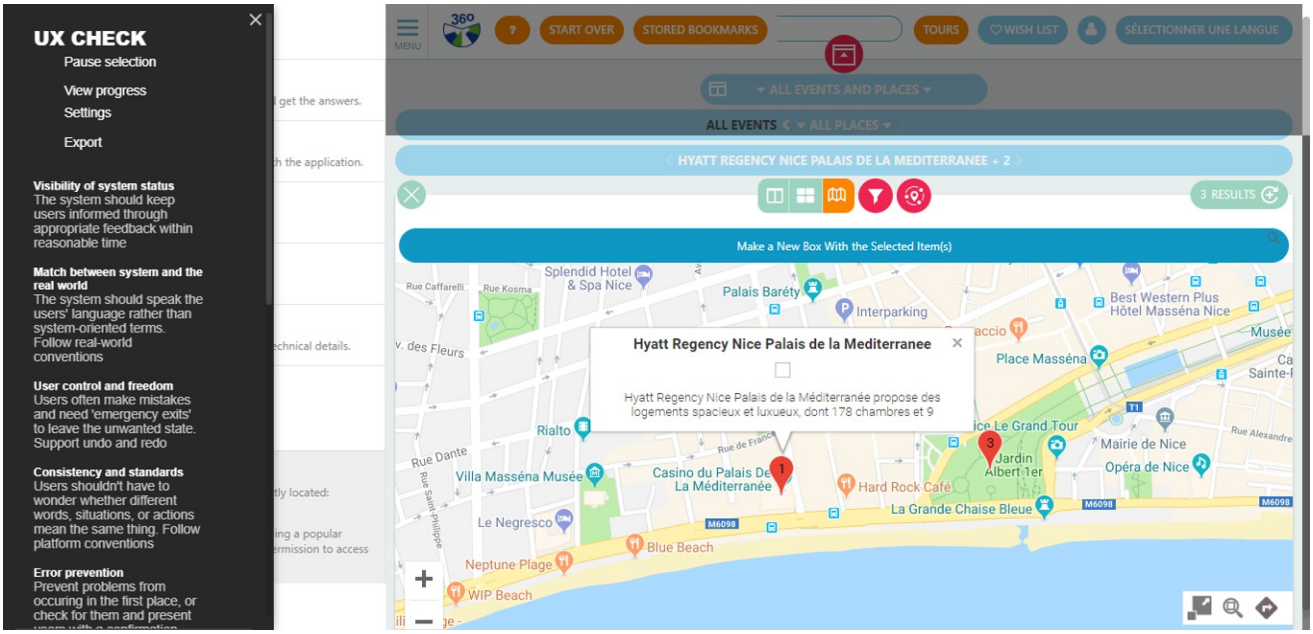


Fig. 6.1: An example on 3cixty of UX Check interface with the impossibility to select a smaller zone on the map.

highlighted element with a red oblong and, as with UX Check, the evaluators can annotate each screenshot with similar indication:

1. title
2. the involved heuristic or criteria selection
3. description
4. recommendation
5. tags
6. severity

Unlike UX CheCk, the screenshots' annotations can always be modified after their validation. Capian is the most advanced tool but, it is a commercial tool and for that reason we cannot choose it.

We want to design a tool in line with these two plugins which includes our heuristics of exploratory search.

6.3 CheXplore plugin design

In this section we introduce **CheXplore**, our Google Chrome plugin which fulfills our specifications in regards of our context of exploratory search. It is available on the Chrome store here: `https://goo.gl/NoA4eG`

We first present the user-centered approach we used to design the plugin and then the resulting plugin through a scenario

6.3.1 Design approach

The design team was composed of a UX researcher (the designer of the model-based heuristics) and of a junior developer, Jean-Marie Dormoy. The user researcher specified the kind of wanted plugin and supervised the design. The developer participated to the design of the plugin and developed it. UX researcher and developer worked very closely.

At each stage of improvement, even the smaller one, CheXplore was tested iteratively by the two design team members. These tests allowed the identifications of issues or possible next improvement. We adopted a user-centric approach: we tested the plugin in real conditions on an exploratory search system and identify the plugin's features issues. For each issues or missing features we seek solutions. We wanted an easy use of the plugin. Table 6.1 shows some examples of the identified issues and the provided solutions.

CheXplore's computer code was based on UX Check's one. Indeed the licence GNU GPL allow us to reuse the part of the UX Check's code. Then, CheXplore was designed with programming languages for the Web: HTML, CSS and Javascript. It was designed on Fedora, a Linux distribution, but the designer of the heuristics uses Windows, so the plugin was tested on two different operating systems and monitor resolutions (1368 x 768 and 1366 x 768).

We want to design a Google Chrome plugin in line with UX Check and Capian. We wanted CheXplore mostly similar to their characteristics:

- the highlighted element on the screenshot must not depend on the HTML tags. Indeed the evaluators have to select the interesting area without any difficulties;
- the URL link is provided for each screenshot;
- the annotation of the screenshot must allow the possibility to:
  - specify the involved heuristic or criteria, and the severity of the issue
  - add comments and recommendation about the identified issue/element
- the evaluation (screenshots and annotation) can be exported in a Word or PDF format. Each document shows: the screenshots (1), the associated annotation (2) and URL link (3)

Identified issues	Provided solution
Issue #1: In the lateral bar, the list of Bastien and Scapin's ergonomic criteria [5] and our heuristics of exploratory search [55] are not visible as a whole. The user needs to scroll.	We decide to divide the criteria and the heuristics in a drop-down menu in terms of the *main criteria* for Bastien and Scapin and *screens* for our heuristics of exploratory search. (e.g. Figure 6.4)
Issue #2: With the same set of criteria and heuristics, the title and definition of each one takes too much space in the lateral bar, even if their are classified by *main criteria* or *screens*: the user has to scroll.	In an evaluation, the user does not need to have the criteria or heuristics' definition always visible in the lateral bar. We decide to add an " ***i*** " button for each of criteria/heuristic which has a definition or description. When the user wants to have more information about one of the criteria or heuristics, she can click on the " ***i*** " button and a little pop-up window appears which contains the associated definition. (e.g. Figure 6.5)
Issue #3: The text font, size and style were not fixe, they followed the website's ones. And sometimes, the size was really small.	We fixed the text front, size and style (Arial) for each element of the Plugin.
Issue #4: The elements saved in View Progress (Figure 6.13) were not emptied out automatically. The screenshots, the annotations and the URL link were gathered and there were not distinction between two evaluation session.	We decide to empty the View Progress list each time the user clicks on "*Start a new evaluation*", i.e. at the beginning of each new evaluation, the view progress empties. This choice allows the possibility to make a break in the evaluation and, even if the user closes the tab, the View Progress list is saved until the user *starts a new evaluation*.
Issue #5: When a user selects a zone on the screenshot and starts to write her annotations, she may be unsatisfied. The only option was to cancel the annotation and take a new screenshot.	We add an icon on the annotation window which allows retaking screenshot (e.g. Figure 6.11). This feature saves the written annotations between two screenshots. Thus, the user does not have to re-write the annotations.

Tab. 6.1: Examples of identified issues and solutions we provided in the adopted iterative design approach of the plugin CheXpole

- the proposed heuristics are: our heuristics of exploratory search, and usability heuristics or criteria (Nielsen's and Bastien and Scapin's ones [47, 5]).

We add to our heuristics of exploratory search two usability heuristics/criteria. Indeed, we showed in Sections 3 and 4 that usability issues can be responsible of exploration difficulties. We want to give the possibility to perform an complete inspection of an interface. The usability is really important for every digital system which involves a human computer interaction. A good usability is also necessary for exploratory search systems, because if an system supports exploratory search process and have a lot of usability issues, the users will have at least the same difficulties to perform an exploratory search task. She can finally find it unusable and useless. Evaluate the usability and exploration aspects are really important and complementary. That is why, on CheXplore, evaluators have the possibility to evaluate also the usability of their exploratory search system.

6.3.2 Scenario-based presentation of CheXplore

We present the plugin we design in a real evaluation task scenario. A fictive evaluator inspects the exploratory search system Frankenplace. We give screenshots in order to illustrate that scenario.

At the beginning of the evaluation, the evaluator is on Frankenplace's home page and open CheXplore (Figure 6.2).

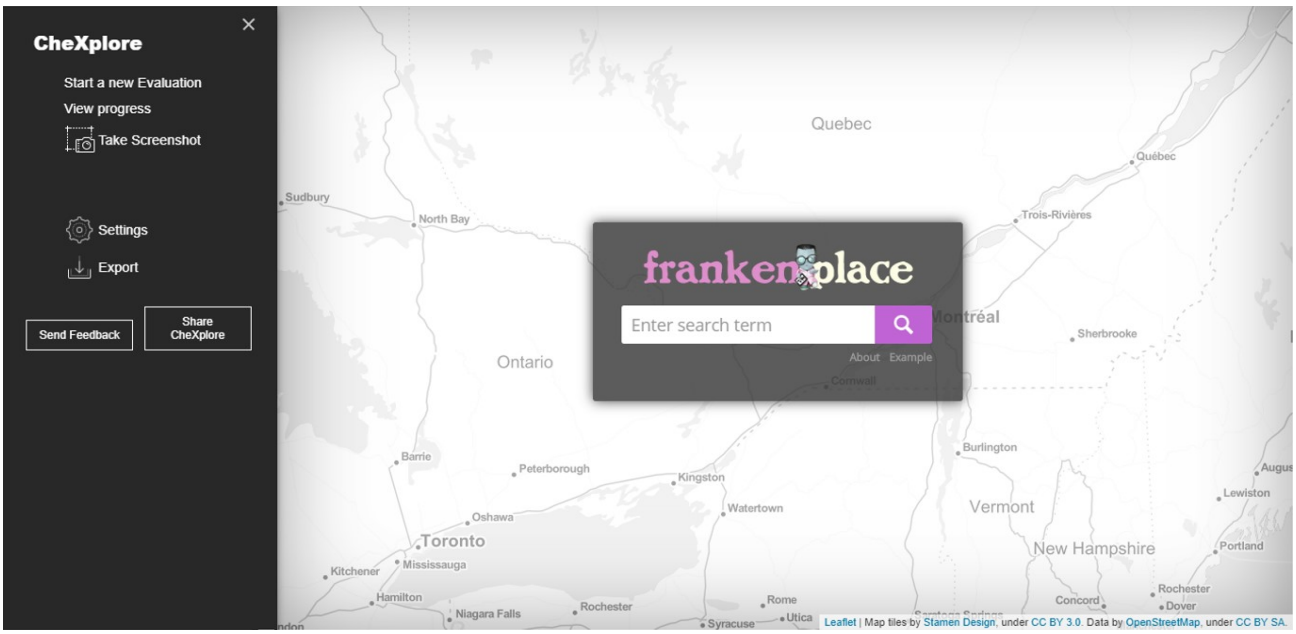


Fig. 6.2: CheXplore - Home page

She clicks on "*start a new evaluation*" on the plugin's lateral bar. She clicks on the drop-down menu to select the *heuristics of exploratory search* (Figure 6.3). She chooses the interrogative form, reads briefly the heuristics and validates the selected set of heuristics.

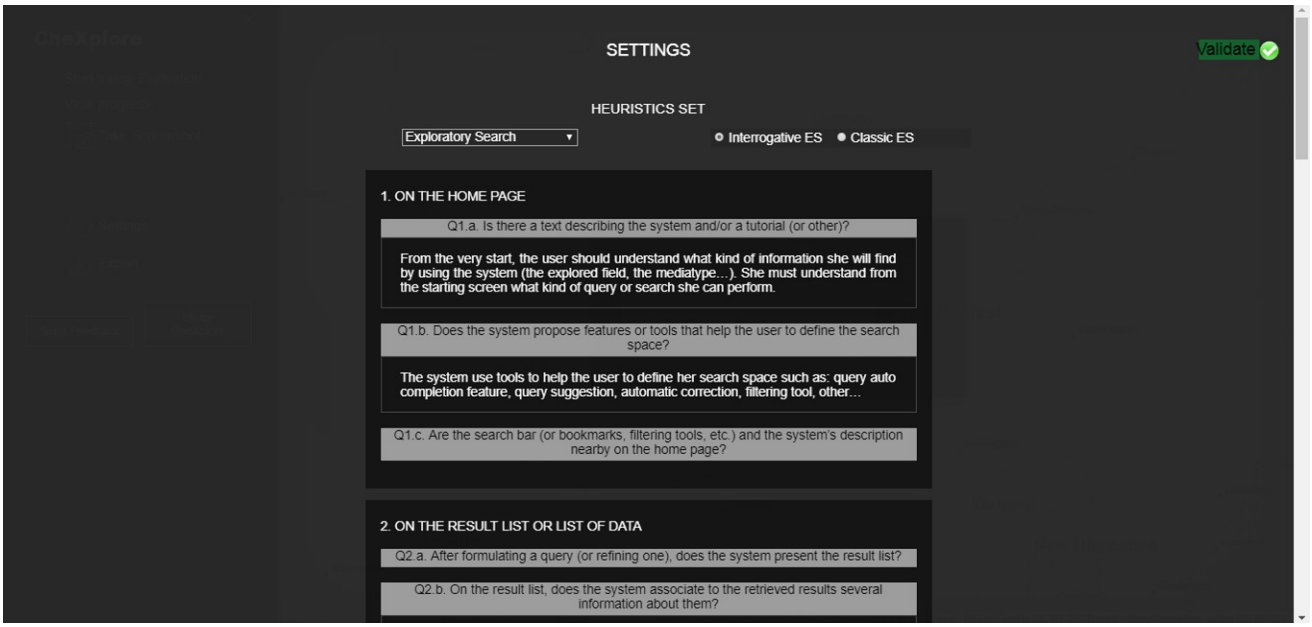


Fig. 6.3: CheXplore - Start an evaluation

Then she selects on the lateral bar the sub-heuristics linked to the home page : she selects on the drop-down menu "*On the home page*" (Figure 6.4).

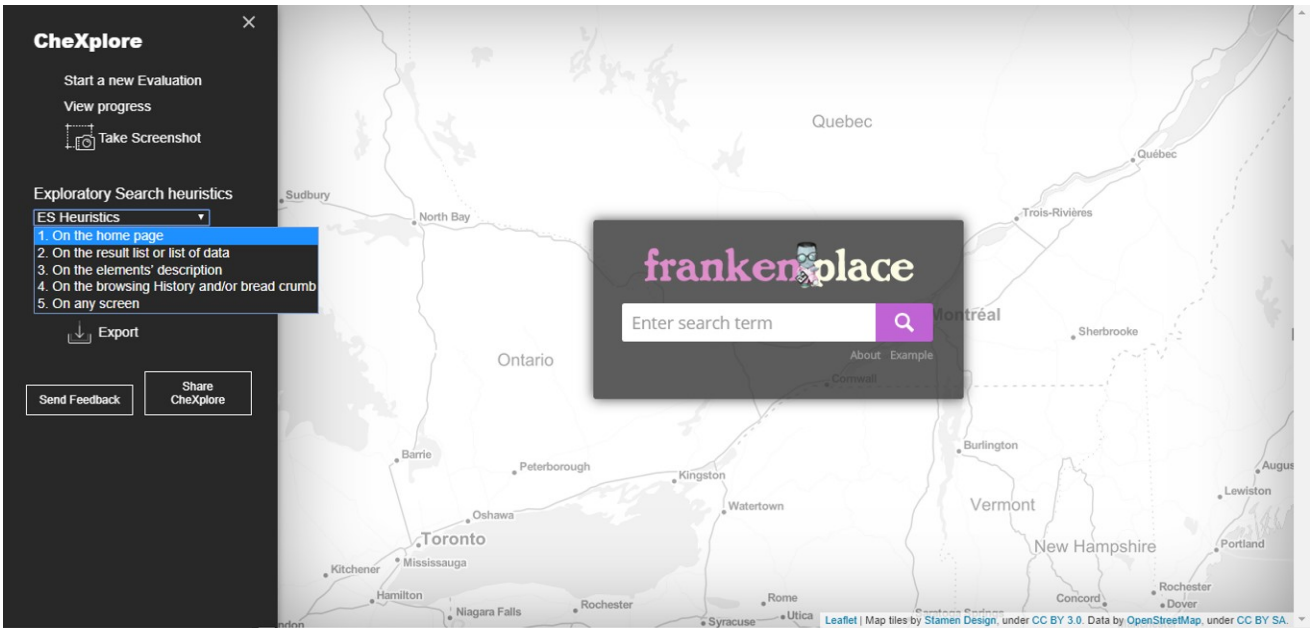


Fig. 6.4: CheXplore - Heuristics selection

She reads the three questions on the lateral bar and clicks on the associated " *i* " button to have more information about the question (Figure 6.5) and closes it.

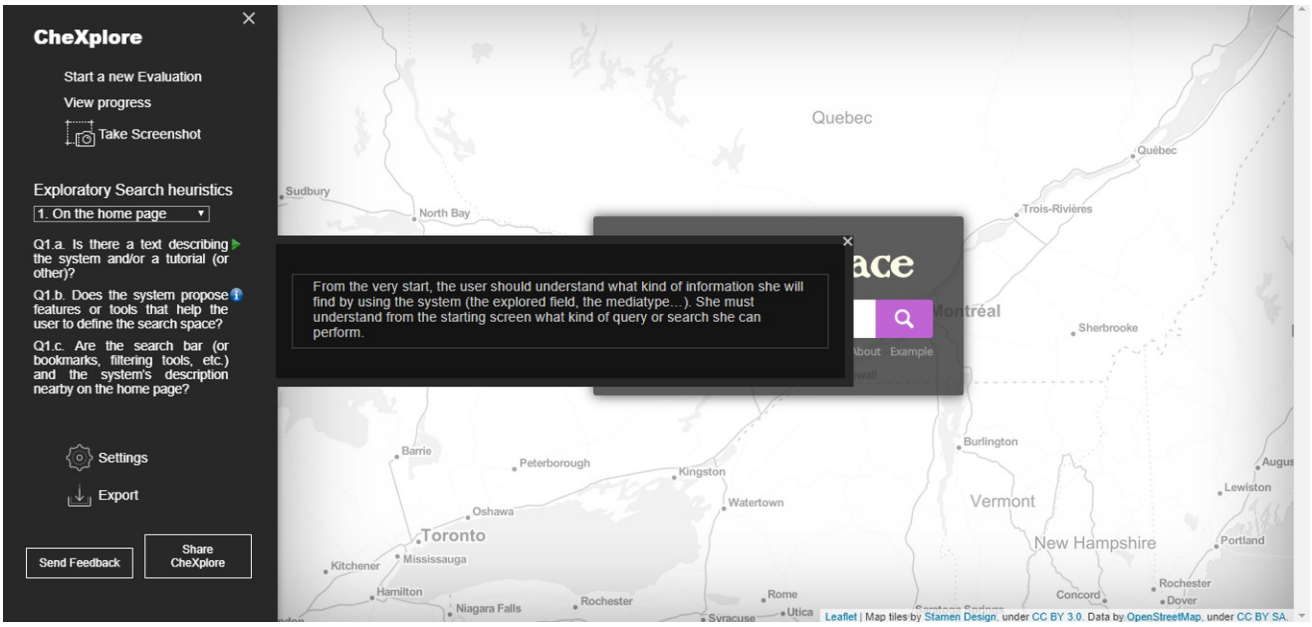


Fig. 6.5: CheXplore - More description button : " *i* "

Then she clicks on the first question (1.a) (Figure 6.6) and defines with her mouse, quite similar as a drag and drop action, the corresponding area on the interface.

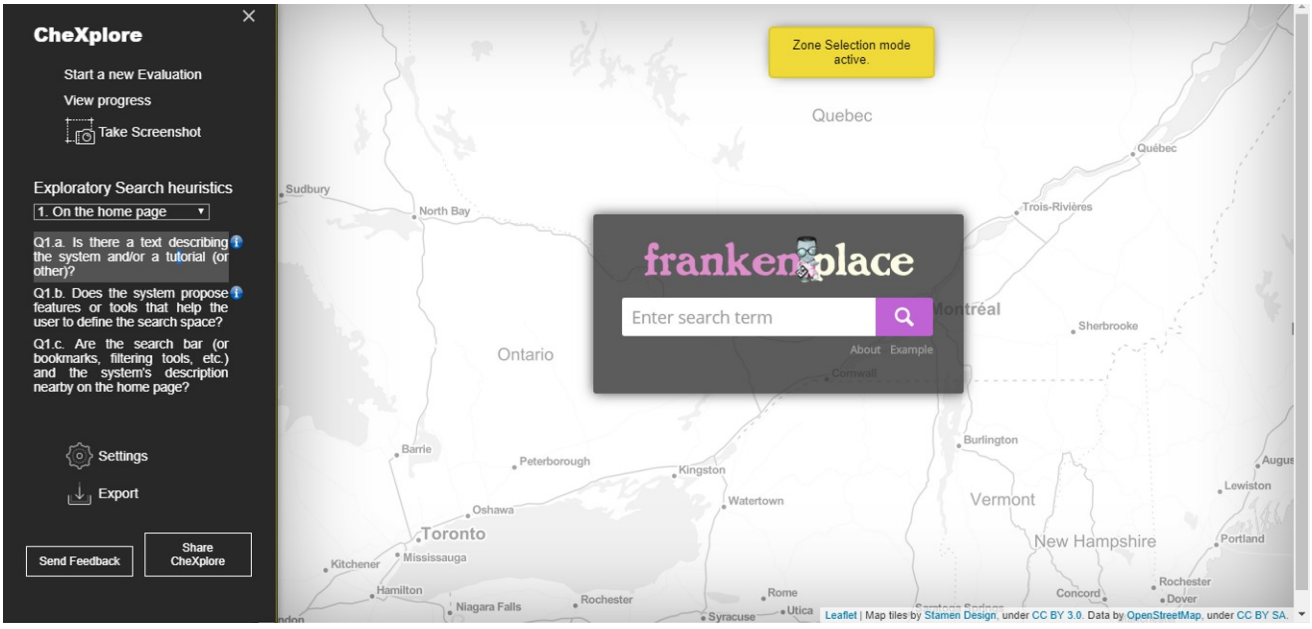


Fig. 6.6: CheXplore - Heuristic selection on the lateral bar before a screenshot

A annotation window appears, and the evaluator decides to move it (Figure 6.7) by clicking on the top of the window.

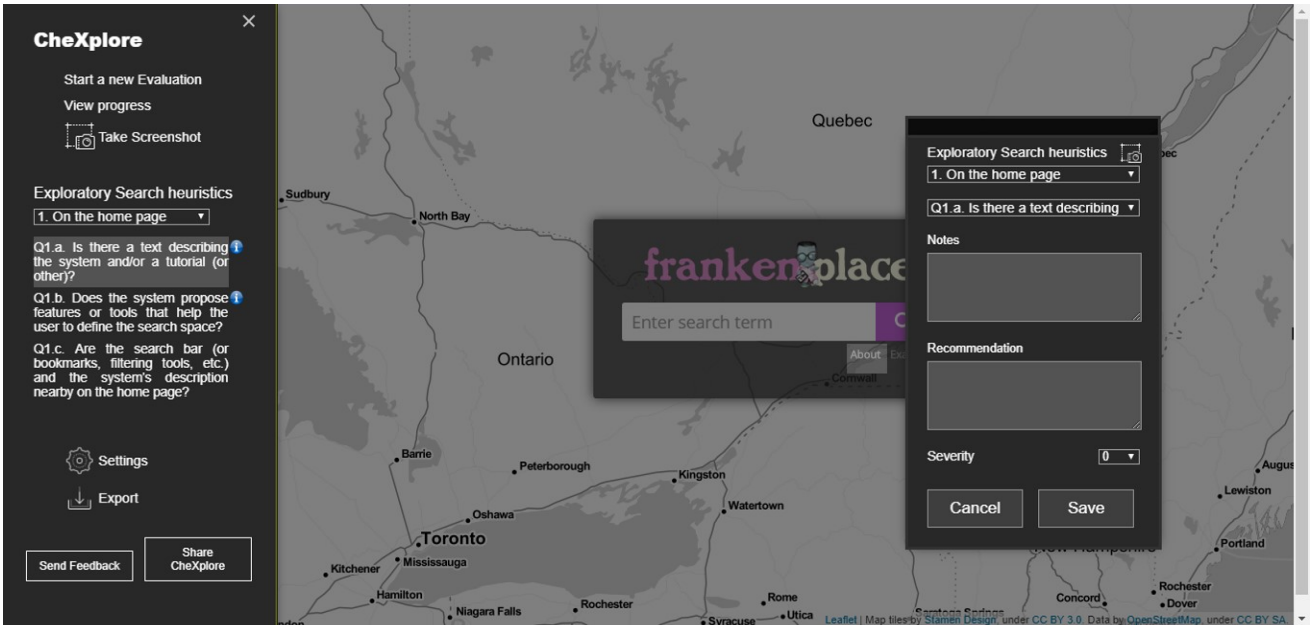


Fig. 6.7: CheXplore - Move of the annotation window

Then, she writes her answer to the question in *note* and recommendation of improvement, if there is one, in *recommendation* (Figure6.8). She can also indicate a severity of the identified element. After fulfilling the annotation window, she clicks on "*save*".

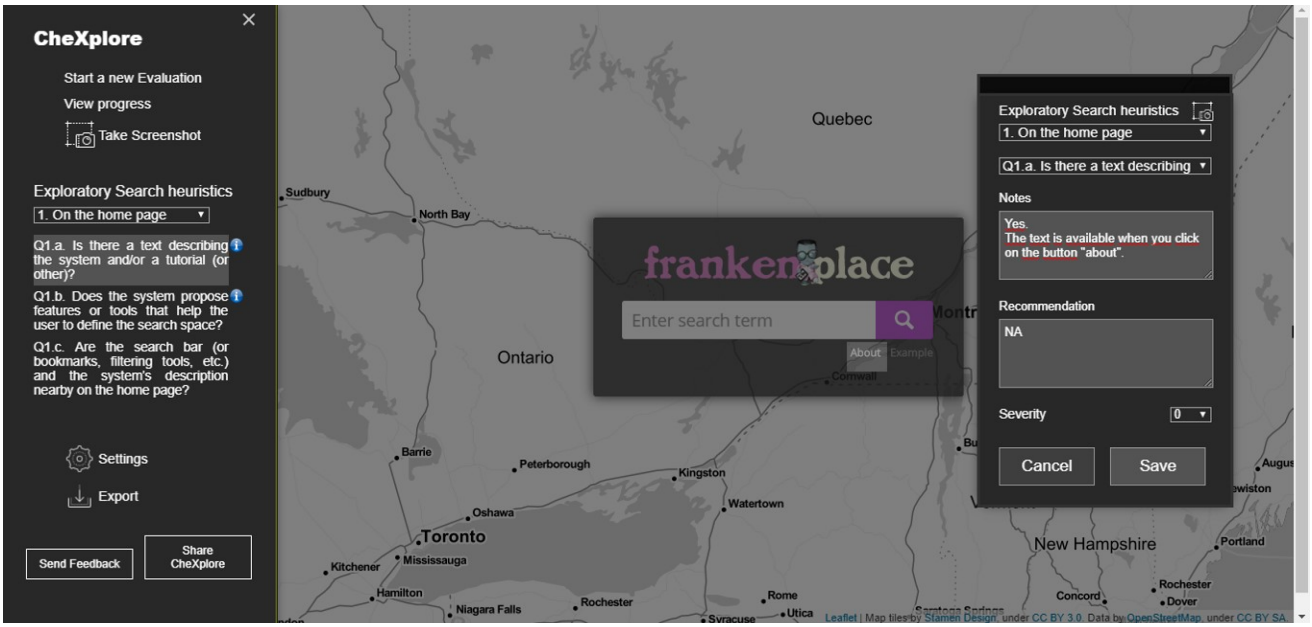


Fig. 6.8: CheXplore - Write the annotations 1

She wants to answer to question 1.b: Frankenplace does not have auto-completion feature. She clicks on the other way to take a screenshot using the "*Take a screenshot*" button (Figure 6.9).

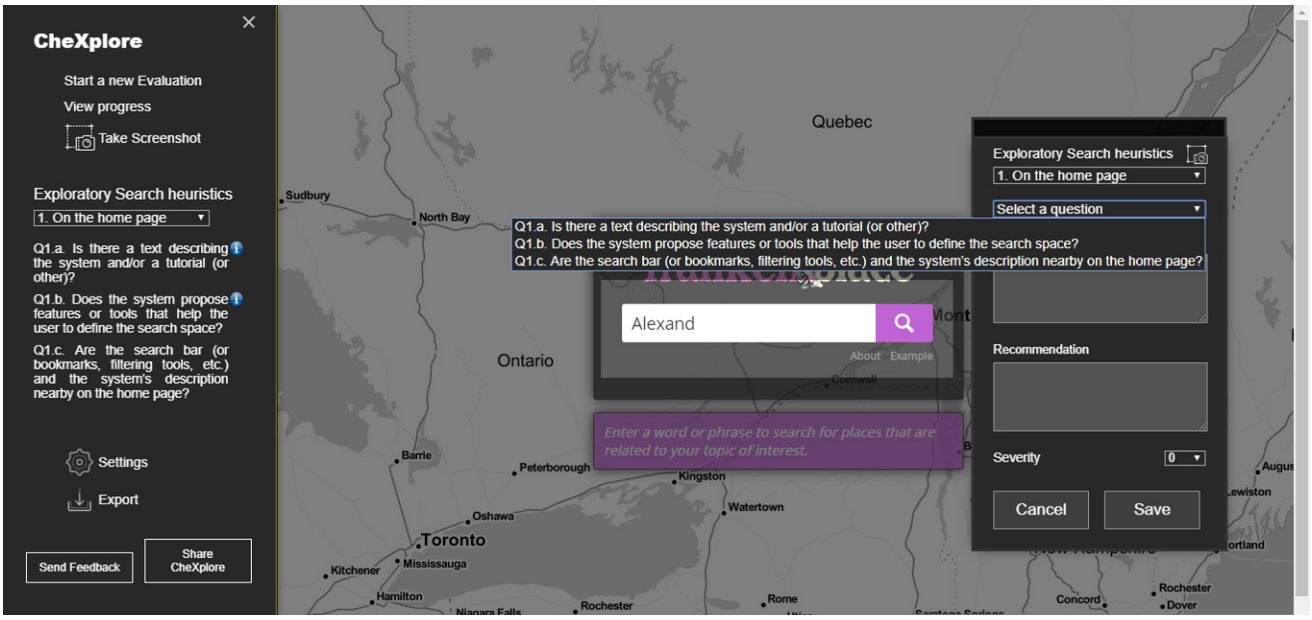


Fig. 6.9: CheXplore - Heuristic selection on the annotation window

She starts to write in the annotation window but she is not satisfied by the first screenshot (Figure 6.10).

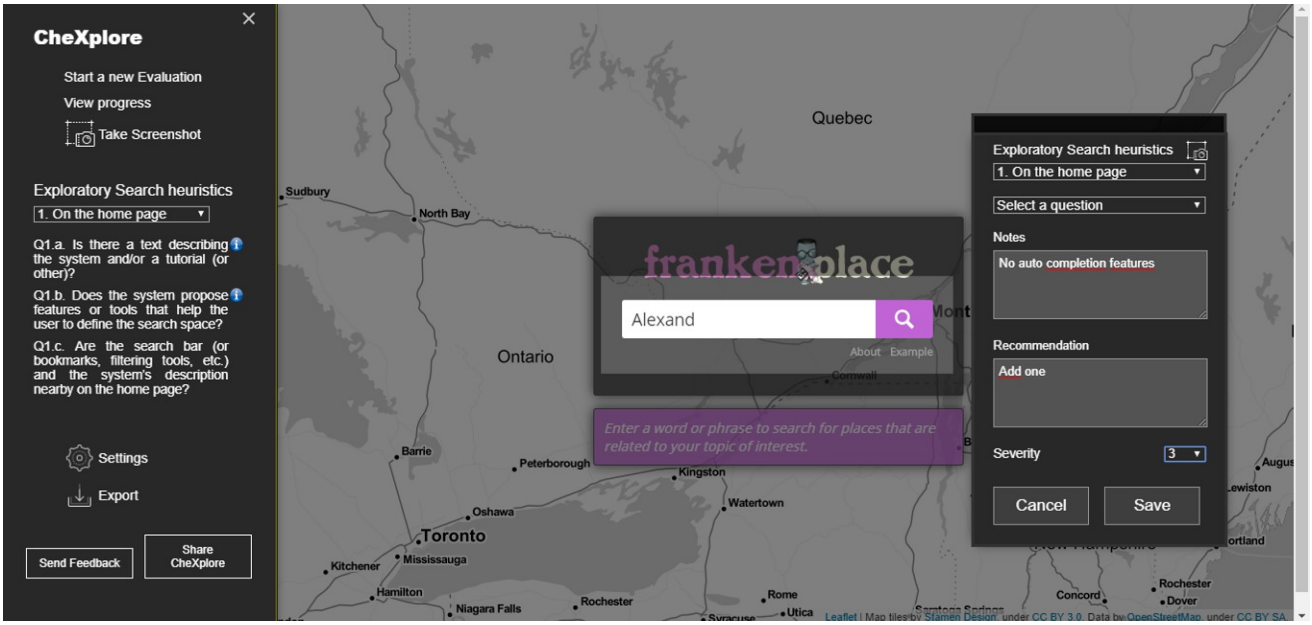


Fig. 6.10: CheXplore - Write the annotations 2

She clicks on the button "*Retake a screenshot*". Between two "*retake a screenshot*" the plugin save the evaluator's annotation. After that, she chooses on the drop-down menu the corresponding heuristic (Figure 6.11).

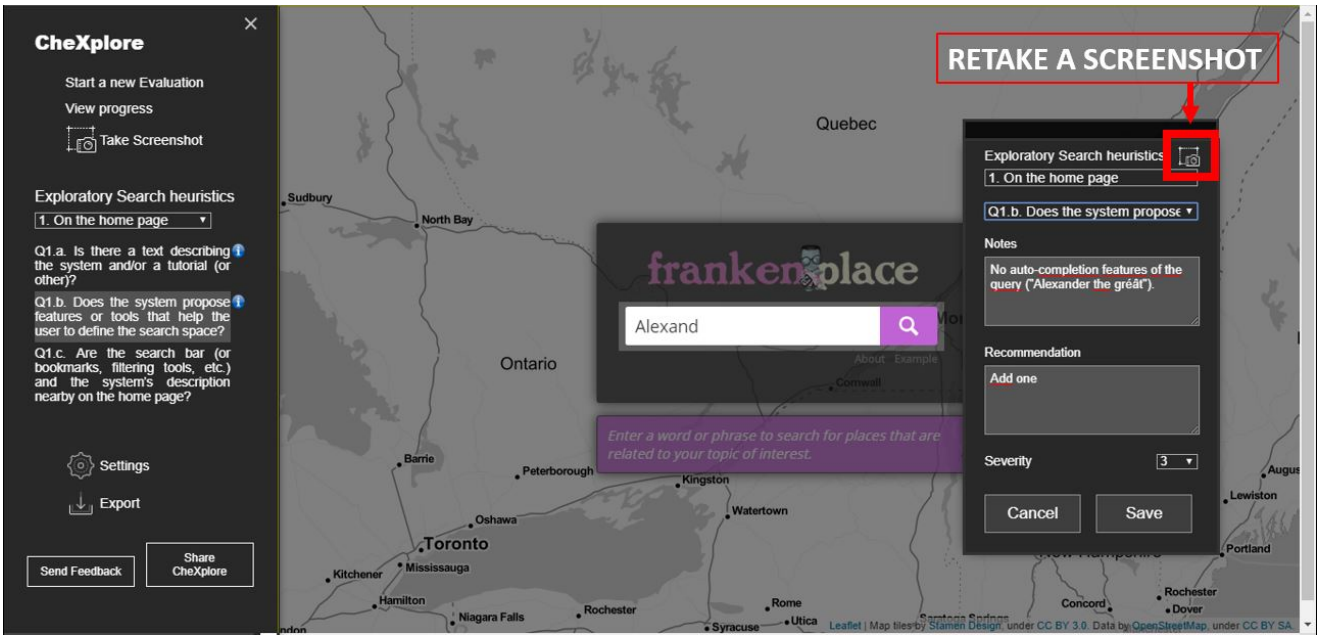


Fig. 6.11: CheXplore - Retake a screenshot with the annotations already written

After answering to question 1.c., the evaluator launches the query "Alexander the great" on Frankenplace. On the drop-down menu of plugin's lateral bar she selects "*On the result list or list of data*" (Figure 6.12) and continues her evaluation.

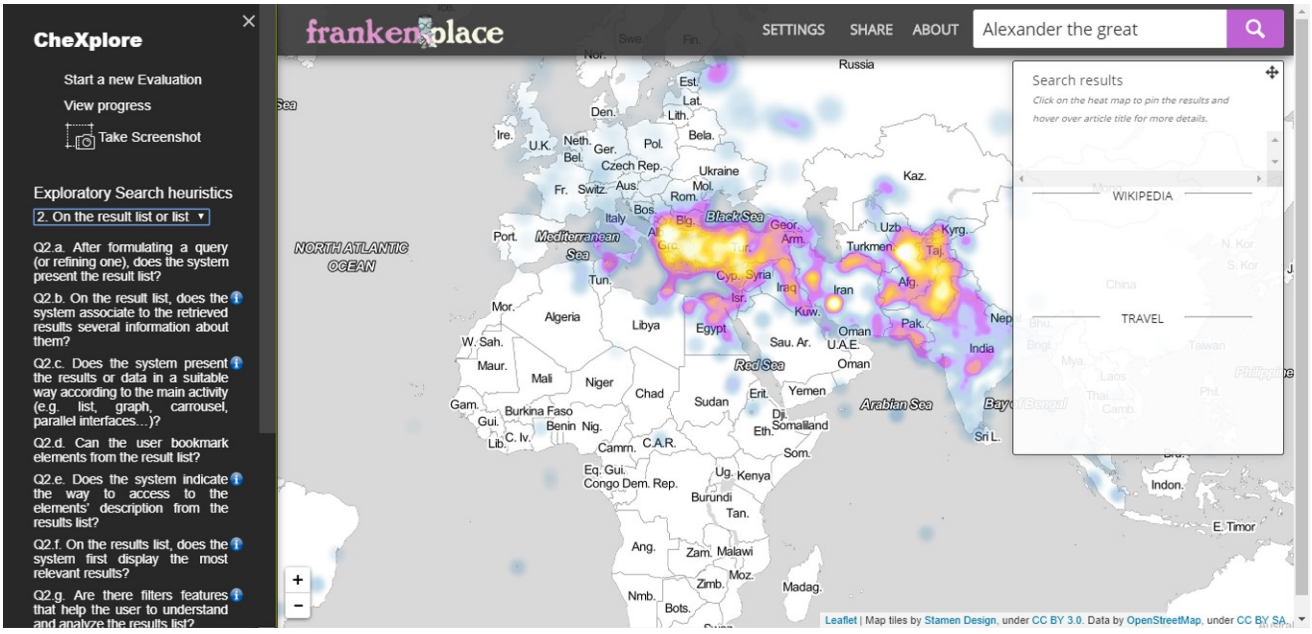


Fig. 6.12: CheXplore - Select another sub-set of heuristics: On the result list of list of data

At one time, she wants to have an overview of the screenshots and the associated annotations: she clicks on "*view progress*" button, observes the saved screenshots and annotations (Figure 6.13), and deletes the last screenshot she took.

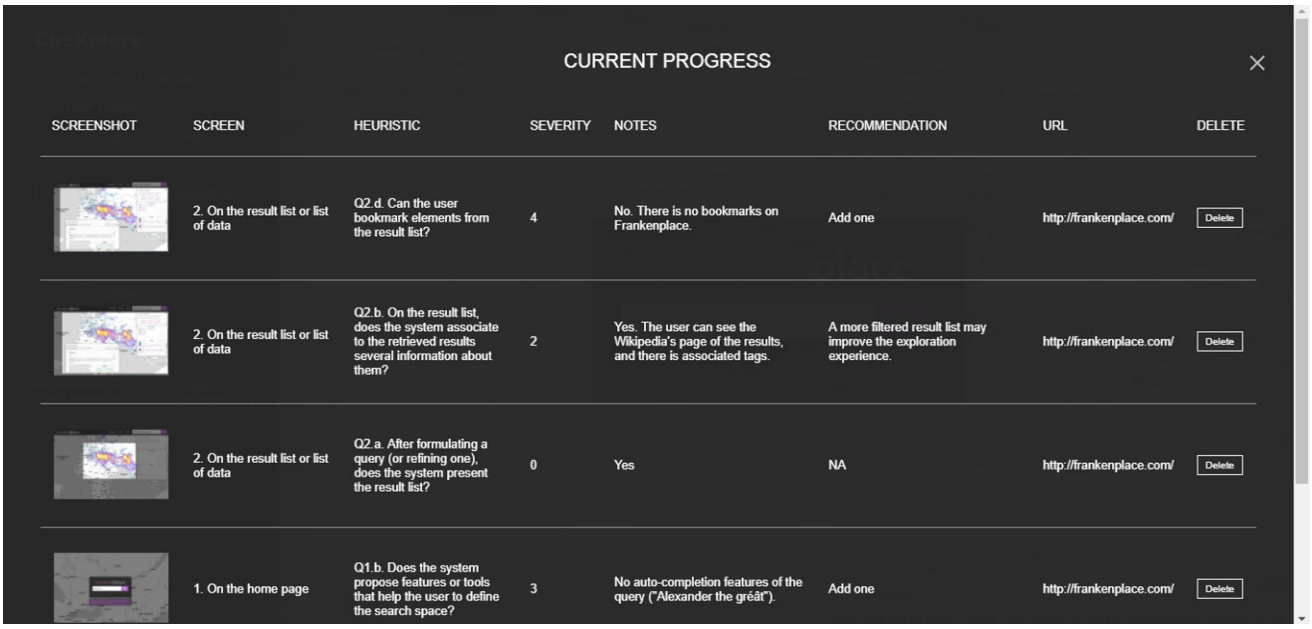


Fig. 6.13: CheXplore - View progress

She stops her evaluation and continues it few hours later: when she opens Frankenplace and the plugin, the screenshots and the annotations were saved and accessible in *View progress*.

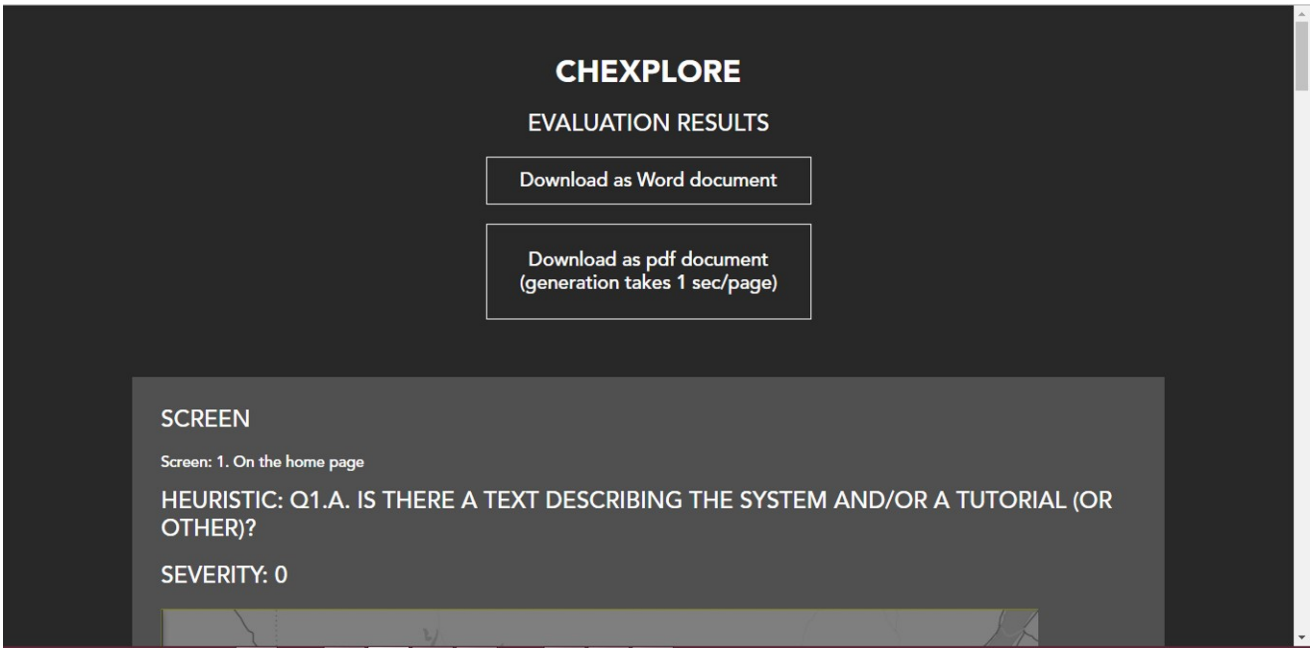


Fig. 6.14: CheXplore - Export a Word or PDF document

At the end of her evaluation she clicks on the "*export*" button on the lateral bar of the plugin(Figure 6.14). A new tab appears, on which she can select the kind of document she wants (Word or PDF) and she can look at her screenshots a last time before the export. She clicks on "*download as a Word document*".

6.4 Conclusion and perspective

In this section we introduced our Google Chrome plugin CheXplore[5]. We adopted user-centric design approach and we presented the plugin with a usage scenario. CheXplore surpasses the limits of the evaluation checklist introduced in Chapter 4 by offering more support in the evaluators' inspection tasks in term of experience and usability of the heuristics of exploratory search. This version is the first functional and usable version of CheXplore which are in line with the specifications we desired. However, it needs some improvements, in terms of usability and code enhancements, for example we already identify some of them:

- Use *CSS Media Queries* for smaller monitor resolution;
- Use less jQuery in the code and more *Vanilla* JavaScript because CheXplore does not works on the exploratory search system Discovery Hub[6]. It is probably linked to the used language because we observed that UX ChecK does not work either on this exploratory search system.
- Add the possibility to add customized heuristics or criteria;
- Add the possibility to remove the lateral bar if preferred;
- Add the possibility to share the evaluation;
- Add the possibility to collaborate on several evaluators' evaluation of a same team.

Some usability tests of the plugin were performed by the UX researcher throughout the design process. Nevertheless, this version of CheXplore needs some evaluation with end-users' feedbacks. Ideally, it would be an evaluation of an exploratory search system with several evaluators involved in the design process of the system. In addition, if we want to identify plugin's usability issues, we can ask to UX designers or UX researcher familiar with heuristics evaluation to test the plugin only with the heuristics and criteria about usability aspects to have a large panel of feedbacks.

[5] https://goo.gl/NoA4eG
[6] `http://discoveryhub.co/`

7 Conclusion

> *What we call the beginning is often the end. And to make an end is to make a beginning. The end is where we start from.*
>
> **— T.S. Eliot**

This chapter summarizes our main contributions to the design of user-centered evaluation methods of exploratory search systems - methods that are based on a model of the exploratory search process and aimed at designers. This chapter also presents some short-term and long-term perspectives.

7.1 Contributions

Our contributions were aimed to overcome the limits of the existing methods of evaluation of exploratory search systems (see Chapter 2). We will present them by indicating to which research question these contributions answered and the chapter where the contributions are detailed. Let us first remind the research questions:

1. Which model of exploratory search process to choose or to elaborate for designing more appropriate evaluation methods?
2. How to elaborate a model of exploratory search process which will be used as a basis for our methods?
3. How to design model-based methods?
4. How to design designer-oriented methods?

First contribution: an approximate model of exploratory search (research questions 1 and 2; Chapter 3) which is used as a basis for the targeted evaluation methods. We first elaborated the eleven characteristics of exploratory search in order to stabilize the literature's description of exploratory search (Section 3.2). These characteristics were used as an analysis grid for the evaluation of several information seeking models. They highlighted that none of the models checked all the characteristics and we had to adapt one of them to the exploratory search process. Thus, we designed our model of exploratory search based on Ellis' model of information

seeking. Our model consists of ten *features* with, as in Ellis' model, a non-linear process and without predefined sequences (Section 3.3). The two-times evaluations validated our model of exploratory search and led us to introduce the notion of *possible transitions* between the model's features (Section 3.4 and 3.5). We thus extended the model by incorporating this notion of *transitions* in it (Sections 3.4.3 and 3.5.3).

Second contribution: a set of heuristics of exploratory search (research questions 3 and 4; Chapter 4). We presented our inspection method of exploratory search systems which consists of a set of heuristics (Section 4.3.2) and a procedure for using them (Section 4.4.2). The heuristics are designed to identify features, interfaces elements or possible actions that support exploratory search behaviors. In the evaluation of the proposed method, we noticed that the interrogative form offer an easier experience of the inspection. In order to support and facilitate the evaluation task we proposed a form format of the heuristics and an online evaluation checklist (Section 4.4.1.2). Indeed, the evaluators, who are computer scientists, need support in their evaluation tasks of exploratory search systems because they are not familiar with inspection methods. The evaluation of the heuristics (with the evaluation checklist) by computer scientists demonstrated that the method significantly help the evaluators identify more and more systematically the presence or the absence of elements or possible actions required to effectively support exploratory search behaviors than without the heuristics (Section 4.5).

Third contribution: elements of a user testing procedure (research questions 3 and 4; Chapter 5). We focused on two model-based elements of a customizable user testing procedure we proposed (Section 5.2): the protocol for the elaboration of exploratory search tasks (Section 5.3) and a video analysis grid for the evaluation of recorded exploratory search sessions associated to users' comments (Section 5.4). The evaluation of the first element, the protocol for the elaboration of exploratory search tasks, showed that all the tasks designed by twenty computer scientists were effectively exploratory search tasks. The first evaluation of the second element, the video analysis grid, aims to assess its capacity to help the participants in the identification of the model's features in video clips. The results showed a large disparity in the scores of the participants and we underscored the possible reasons and factors that can explain such results. From these results analysis, we suggested several improvements to the video analysis grid.

Fourth contribution: a tool for exploratory search interface inspection (research question 4; Chapter 6). We proposed a tool facilitating the evaluation with our heuristics of exploratory search. We introduced our Google Chrome Plugin *CheXplore* and presented the design approach. This plugin aims to address an issue which has been highlighted in Chapter 4: the online evaluation checklist does not

completely ease the use of the heuristics of exploratory search. Indeed, it implies a split-screen display, or if it not possible several switching between two browser tabs (the evaluated exploratory search system and the evaluation checklist). CheXplore allows the evaluators to inspect the interface, by analyzing it in terms of heuristics of exploratory search, and also in terms of usability heuristics (Bastien and Scapin; Nielsen). The results of the analysis (including screenshots and comments) are recorded and are included in an analysis report that can be printed.

7.2 Perspectives

The work presented in this thesis is at the intersection of different research domains: cognitive ergonomics, human computer interaction, cognitive psychology, information seeking and retrieval, computer sciences (*knowledge engineering*). Consequently, the inderdisciplinary nature of the contributions can be seen as valuable, at different levels, for each of these research domains. In addition to designers of exploratory search systems, anyone interesting in exploratory search concept or system could benefit from the contributions exposed in this thesis. The model of exploratory search can also be a contribution to information seeking field. Researchers or practitioners in human-computer interaction interested in the elaboration of new evaluation methods of prospective human-computer interactions (*e.g.* computer scientists, cognitive ergonomists, UX researchers...) could be concerned by all the contributions exposed here.

We present the perspectives in two subsections below. The first one presents the short-term improvements which are in the wake of our contributions. The second one opens new perspectives in the long-run of possible other ways of evaluating of exploratory search inspired on our work.

7.2.1 Short-term perspectives

In the short-term, we propose to improve and further evaluate the work presented in this thesis:

1. The heuristics of exploratory search need further evaluation: the method, and the heuristics in their form format, were not tested in real conditions with a real team of computer scientists who inspect individually their system and confront their results. Several evaluations involving such participants would highlight the difficulty to understand the *inspection method* process. In addition, these evaluations would enable us to design a more accessible procedure to use the heuristics for non familiar evaluators;

2. The plugin CheXplore needs to be user tested to get end-users' feedbacks. Ideally, this evaluation could be included in the evaluation of the heuristics previously mentioned. Furthermore, it would help us identify usability issues of the plugin which we will add to the already suggested improvements.
3. The video analysis grid needs further evaluation and improvement to evaluate its understandability and its relevance. Moreover, a tool facilitating the video annotation task with our grid would be extremely helpful for the users.
4. The user testing method needs to be evaluated in real conditions, with the same team of computer scientists mentioned earlier for example. This would help us to evaluate its understandability and to improve it accordingly.

7.2.2 Long-term perspectives

We propose in this thesis two evaluation methods of exploratory search which are focused on users' behavior. However, the evaluation of exploratory search systems is still an open issue, and it does remain other interesting ways of evaluating them. The long-term perspectives we envision are related to them.

Indeed, in Chapter 3, we mentioned a work we did on the evaluation of algorithms with respect to the relevance of exploratory search results [39]. It was an attempt of the analysis of exploratory search results relevance, but this work can be extended by establishing a list of characteristics of exploratory search results which supports the knowledge acquisition and allows serendipitous discoveries. We do believe that the results have to be varied to offer the user a large panel of possible behaviors. Exploratory search systems have to offer a subtle mix of results highly relevant to the goal at hand, and of not directly linked results which can be surprising and interesting. This corresponds to a first new research question : how to assess the quality of the results in an exploratory search?

More broadly, we can also create *criteria of explorability* which could be used on any search systems which also support a more important part of *lookup activities* than exploratoy search systems. Indeed, exploratory search can start from a lookup search, and these *criteria of explorability*, could explain, for example, how to encourage and support exploratory search on a *classic* search engine (*e.g.* Google, Yahoo!, Qwant, Bing...), or in any system or website with whom the users can explore information. These criteria could be inspired in large part from the heuristics of exploratory search, avoiding the ones which are specific to exploratory search systems. This corresponds to a second new research question: how to induce and support exploratory search in any kind of search system?

Bibliography

[1] Benjamin Adams and Grant McKenzie. „Frankenplace: An Application for Similarity-Based Place Search.“ In: *ICWSM*. 2012 (cit. on pp. 12, 13).

[2] Jae-Wook Ahn and Peter Brusilovsky. „Adaptive visualization for exploratory information retrieval“. In: *Information Processing & Management* 49.5 (2013), pp. 1139–1164 (cit. on pp. 4, 14).

[3] Paul André, Jaime Teevan, and Susan T Dumais. „From x-rays to silly putty via Uranus: serendipity and its role in web search“. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM. 2009, pp. 2033–2036 (cit. on p. 23).

[4] Kumaripaba Athukorala, Dorota Głowacka, Giulio Jacucci, Antti Oulasvirta, and Jilles Vreeken. „Is exploratory search different? A comparison of information search behavior for exploratory and lookup tasks“. In: *Journal of the Association for Information Science and Technology* 67.11 (2016), pp. 2635–2651 (cit. on pp. 21, 22, 31, 77).

[5] JM Christian Bastien and Dominique L Scapin. „Ergonomic criteria for the evaluation of human-computer interfaces“. PhD thesis. Inria, 1993 (cit. on pp. 54, 71, 72, 106, 109, 110).

[6] Marcia J Bates. „The design of browsing and berrypicking techniques for the online search interface“. In: *Online review* 13.5 (1989), pp. 407–424 (cit. on pp. 25, 27, 28).

[7] Marcia J Bates. „Where should the person stop and the information search interface start?“ In: *Information Processing & Management* 26.5 (1990), pp. 575–591 (cit. on p. 14).

[8] Nicholas J Belkin. „Anomalous states of knowledge as a basis for information retrieval“. In: *Canadian journal of information science* 5.1 (1980), pp. 133–143 (cit. on p. 7).

[9] Nicholas J. Belkin, Pier Giorgio Marchetti, and Colleen Cool. „BRAQUE: Design of an interface to support user interaction in information retrieval“. In: *Information processing & management* 29.3 (1993), pp. 325–344 (cit. on p. 14).

[10] Benjamin S Bloom, Max D Engelhart, Edward J Furst, Walker H Hill, and David R Krathwohl. *Taxonomy of educational objectives, handbook I: The cognitive domain*. Vol. 19. New York: David McKay Co Inc, 1956 (cit. on pp. 31, 79).

[11] Nicole Boubée and André Tricot. *Qu'est-ce que rechercher de l'information?: état de l'art*. Presses de l'Enssib, 2010 (cit. on pp. 27, 28, 30, 34).

[12]Alessandro Bozzon, Marco Brambilla, Stefano Ceri, and Davide Mazza. „Exploratory search framework for Web data sources“. In: *The VLDB Journal* 22.5 (2013), pp. 641–663 (cit. on pp. 4, 13, 14).

[13]Silvana Castano, Alfio Ferrara, and Stefano Montanelli. „inWalk: Interactive and Thematic Walks inside the Web of Data.“ In: *EDBT*. Citeseer. 2014, pp. 628–631 (cit. on pp. 11, 13).

[14]Sylvie Catellin. *Sérendipité. Du conte au concept: Du conte au concept*. Le Seuil, 2014 (cit. on pp. 23, 24).

[15]Sylvie Catellin and Laurent Loty. „Sérendipité et indisciplinarité“. In: *Hermès, La Revue* 3 (2013), pp. 32–40 (cit. on p. 23).

[16]Gilbert Cockton. „Usability evaluation“. In: (2012) (cit. on p. 53).

[17]David Ellis. „A behavioural approach to information retrieval system design“. In: *Journal of documentation* 45.3 (1989), pp. 171–212 (cit. on pp. 25, 26).

[18]David Ellis, Deborah Cox, and Katherine Hall. „A comparison of the information seeking patterns of researchers in the physical and social sciences“. In: *Journal of documentation* 49.4 (1993), pp. 356–369 (cit. on pp. 25, 26).

[19]Allen Foster and Nigel Ford. „Serendipity and information seeking: an empirical study“. In: *Journal of documentation* 59.3 (2003), pp. 321–340 (cit. on p. 23).

[20]Dorota Glowacka, Evangelos Milios, Axel J Soto, and Fernando Paulovich. „Exploratory Search and Interactive Data Analytics“. In: *Proceedings of the 22nd International Conference on Intelligent User Interfaces Companion*. ACM. 2017, pp. 9–11 (cit. on p. 2).

[21]Dorota Głowacka, Evangelos Milios, Axel J Soto, Fernando V Paulovich, and Denis Parra. „Second Workshop on Exploratory Search and Interactive Data Analytics (ESIDA)“. In: (2018) (cit. on p. 2).

[22]Gene Golovchinsky, John Adcock, Jeremy Pickens, Pernilla Qvarfordt, and Maribeth Back. „Cerchiamo: a collaborative exploratory search tool“. In: *Proceedings of Computer Supported Cooperative Work (CSCW)* (2008), pp. 8–12 (cit. on pp. 12, 13).

[23]Ahmed Hassan Awadallah, Ryen W White, Patrick Pantel, Susan T Dumais, and Yi-Min Wang. „Supporting complex search tasks“. In: *Proceedings of the 23rd ACM International Conference on Conference on Information and Knowledge Management*. ACM. 2014, pp. 829–838 (cit. on p. 23).

[24]Chathra Hendahewa and Chirag Shah. „Implicit search feature based approach to assist users in exploratory search tasks“. In: *Information Processing & Management* 51.5 (2015), pp. 643–661 (cit. on p. 23).

[25]Peter Ingwersen and Kalervo Järvelin. *The turn: Integration of information seeking and retrieval in context*. Vol. 18. Springer Science & Business Media, 2006 (cit. on pp. 26, 28).

[26]Diane Kelly, Susan Dumais, and Jan O Pedersen. „Evaluation challenges and directions for information-seeking support systems“. In: *Computer* 42.3 (2009), pp. 60–66 (cit. on p. 13).

[27]Eunyee Koh, Andruid Kerne, and Rodney Hill. „Creativity support: information discovery and exploratory search“. In: *Proceedings of the 30th annual international ACM SIGIR conference on Research and development in information retrieval*. ACM. 2007, pp. 895–896 (cit. on p. 23).

[28]Carol C Kuhlthau. „Inside the search process: Information seeking from the user's perspective“. In: *Journal of the American society for information science* 42.5 (1991), p. 361 (cit. on pp. 14, 26, 28, 29).

[29]Bill Kules and Robert Capra. „Creating exploratory tasks for a faceted search interface“. In: *Proc. of HCIR 2008* (2008), pp. 18–21 (cit. on p. 78).

[30]Bill Kules and Robert Capra. „Designing exploratory search tasks for user studies of information seeking support systems“. In: *Proceedings of the 9th ACM/IEEE-CS joint conference on Digital libraries*. ACM. 2009, pp. 419–420 (cit. on p. 77).

[31]Bill Kules, Robert Capra, Matthew Banta, and Tito Sierra. „What do exploratory searchers look at in a faceted search interface?“ In: *Proceedings of the 9th ACM/IEEE-CS joint conference on Digital libraries*. ACM. 2009, pp. 313–322 (cit. on p. 78).

[32]Stephann Makri, Ann Blandford, and Anna L Cox. „Using information behaviors to evaluate the functionality and usability of electronic resources: From Ellis's model to evaluation“. In: *Journal of the American Society for Information Science and Technology* 59.14 (2008), pp. 2244–2267 (cit. on p. 34).

[33]Gary Marchionini. „Exploratory search: from finding to understanding“. In: *Communications of the ACM* 49.4 (2006), pp. 41–46 (cit. on pp. 3, 7, 8, 19–21, 26, 30, 31, 79).

[34]Gary Marchionini. „Information seeking in electronic environments“. In: (1995) (cit. on pp. 26, 28–30, 44).

[35]Gary Marchionini and Ryen White. „Find what you need, understand what you find“. In: *International Journal of Human [# x02013] Computer Interaction* 23.3 (2007), pp. 205–237 (cit. on p. 35).

[36]Nicolas Marie. „Linked data based exploratory search“. PhD thesis. Université Nice Sophia Antipolis, 2014 (cit. on pp. 7, 11, 13, 20–23).

[37]Nicolas Marie, Fabien Gandon, Myriam Ribière, and Florentin Rodio. „Discovery hub: on-the-fly linked data exploratory search“. In: *Proceedings of the 9th International Conference on Semantic Systems*. ACM. 2013, pp. 17–24 (cit. on p. 11).

[38]Nicolas Marie, Fabien Gandon, Damien Legrand, and Myriam Ribière. „Exploratory search on the top of dbpedia chapters with the discovery hub application“. In: *Extended Semantic Web Conference*. Springer. 2013, pp. 184–188 (cit. on p. 11).

[39]Nicolas Marie, Fabien Gandon, Alain Giboin, and Émilie Palagi. „Exploratory search on topics through different perspectives with DBpedia“. In: *Proceedings of the 10th International Conference on Semantic Systems*. ACM. 2014, pp. 45–52 (cit. on pp. 11, 20, 21, 23, 121).

[40]Roberto Mirizzi and Tommaso Di Noia. „From exploratory search to web search and back“. In: *Proceedings of the 3rd workshop on Ph. D. students in information and knowledge management*. ACM. 2010, pp. 39–46 (cit. on pp. 11, 13).

[41]Roberto Mirizzi, Azzurra Ragone, Tommaso Di Noia, and Eugenio Di Sciascio. „Semantic wonder cloud: exploratory search in DBpedia“. In: *International Conference on Web Engineering*. Springer. 2010, pp. 138–149 (cit. on pp. 11, 13).

[42]Arunav Mishra and Klaus Berberich. „ExposÉ: Exploring past news for seminal events“. In: *Proceedings of the 24th International Conference on World Wide Web*. ACM. 2015, pp. 223–226 (cit. on pp. 12, 13).

[43]Alberto Musetti, Andrea Giovanni Nuzzolese, Francesco Draicchio, et al. „Aemoo: Exploratory search based on knowledge patterns over the semantic web“. In: *Semantic Web Challenge* 136 (2012) (cit. on pp. 11, 13).

[44]Thi-Nhu Nguyen, Duy-Thanh Dinh, and Tuan-Dung Cao. „Empowering exploratory search on linked movie open data with semantic technologies“. In: *Proceedings of the Sixth International Symposium on Information and Communication Technology*. ACM. 2015, pp. 296–303 (cit. on pp. 12, 13, 20).

[45]Jakob Nielsen. „10 usability heuristics for user interface design“. In: *Nielsen Norman Group* 1.1 (1995) (cit. on p. 54).

[46]Jakob Nielsen. „Enhancing the explanatory power of usability heuristics“. In: *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*. ACM. 1994, pp. 152–158 (cit. on p. 53).

[47]Jakob Nielsen. „Usability inspection methods“. In: *Conference companion on Human factors in computing systems*. ACM. 1994, pp. 413–414 (cit. on pp. 53, 54, 64, 72, 106, 110).

[48]Jakob Nielsen. *Why you only need to test with 5 users*. 2000 (cit. on p. 76).

[49]Jakob Nielsen and Rolf Molich. „Heuristic evaluation of user interfaces“. In: *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM. 1990, pp. 249–256 (cit. on p. 68).

[50]Mark Nolan. „IA column: Exploring exploratory search“. In: *Bulletin of the American Society for Information Science and Technology* 34.4 (2008), pp. 38–41 (cit. on pp. 4, 14).

[51]DA Norman. *Cognitive Engineering.(Ed.), User Centered System Design (pp. 31-61)*. 1986 (cit. on p. 18).

[52]Don Norman. *The design of everyday things: Revised and expanded edition*. Constellation, 2013 (cit. on p. 18).

[53]Andrea Giovanni Nuzzolese, Valentina Presutti, Aldo Gangemi, Alberto Musetti, and Paolo Ciancarini. „Aemoo: exploring knowledge on the web“. In: *Proceedings of the 5th Annual ACM Web Science Conference*. ACM. 2013, pp. 272–275 (cit. on p. 13).

[54]Emilie Palagi. „Quelle méthode ergonomique élaborer pour évaluer les moteurs de recherche exploratoire?“ In: *COnférence en Recherche d'Information et Applications 2015 (CORIA 2015)*. 2015 (cit. on pp. 13, 20–22).

[55]Emilie Palagi, Fabien Gandon, Alain Giboin, and Raphaël Troncy. „A Model-based Heuristic Evaluation Method of Exploratory Search“. In: *HCI 2018-32nd British Human Computer Interaction Conference*. 2018, pp. 1–5 (cit. on pp. 5, 54, 105, 109).

[56]Emilie Palagi, Fabien Gandon, Alain Giboin, and Raphaël Troncy. „A Survey of Definitions and Models of Exploratory Search“. In: *Proceedings of the 2017 ACM Workshop on Exploratory Search and Interactive Data Analytics*. ACM. 2017, pp. 3–8 (cit. on pp. 5, 8, 19, 24, 32, 35).

[57]Emilie Palagi, Alain Giboin, Fabien Gandon, and Raphäel Troncy. „Un Modèle de Recherche Exploratoire pour l'Évaluation de ses Systèmes & Applications“. In: *IHM 2018-30eme conférence francophone sur l'Interaction Homme-Machine*. 2018, pp. 1–7 (cit. on p. 5).

[58]Eli Pariser. *The filter bubble: What the Internet is hiding from you*. Penguin UK, 2011 (cit. on p. 1).

[59]M Cristina Pattuelli, Matt Miller, Leanora Lange, Sean Fitzell, and Carolyn Li-Madeo. „Crafting linked open data for cultural heritage: Mapping and curation tools for the linked jazz project“. In: *Code4Lib Journal* 21 (2013) (cit. on pp. 12, 13).

[60]Peter Pirolli. *Information foraging theory: Adaptive interaction with information*. Oxford University Press, 2007 (cit. on p. 9).

[61]Peter Pirolli and Stuart Card. „Information foraging.“ In: *Psychological review* 106.4 (1999), p. 643 (cit. on p. 8).

[62]Peter Pirolli and Stuart Card. „Information foraging in information access environments“. In: *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM Press/Addison-Wesley Publishing Co. 1995, pp. 51–58 (cit. on p. 8).

[63]Yan Qu and George W Furnas. „What do the attributes of exploratory search tell us about evaluation“. In: *EESS 2006* (2006), p. 51 (cit. on p. 13).

[64]Harald Sack. „The journey is the reward-towards new paradigms in web search“. In: *International Conference on Business Information Systems*. Springer. 2015, pp. 15–26 (cit. on p. 21).

[65]Dominique L Scapin and JM Christian Bastien. „Ergonomic criteria for evaluating the ergonomic quality of interactive systems“. In: *Behaviour & information technology* 16.4-5 (1997), pp. 220–231 (cit. on p. 54).

[66]Georg Singer, Ulrich Norbisrath, Eero Vainikko, Hannu Kikkas, and Dirk Lewandowski. „Search-logger analyzing exploratory search tasks“. In: *Proceedings of the 2011 ACM Symposium on Applied Computing*. ACM. 2011, pp. 751–756 (cit. on p. 77).

[67]Deborah Hinderer Sova, Jakob Nielsen, NN GROUP, et al. „234 tips and tricks for recruiting users as participants in usability studies“. In: *URL: http://www. nngroup. com/reports/tips/recruiting/234_recruiting_tips. pdf* (2003) (cit. on p. 76).

[68]Rick Spencer. „The streamlined cognitive walkthrough method, working around social constraints encountered in a software development company“. In: *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*. ACM. 2000, pp. 353–359 (cit. on p. 54).

[69]Andreas Spitz, Satya Almasian, and Michael Gertz. „EVELIN: Exploration of Event and Entity Links in Implicit Networks“. In: *Proceedings of the 26th International Conference on World Wide Web Companion*. International World Wide Web Conferences Steering Committee. 2017, pp. 273–277 (cit. on pp. 12, 13).

[70]Raphael Troncy, Giuseppe Rizzo, Anthony Jameson, et al. „3cixty: Building comprehensive knowledge bases for city exploration“. In: *Web Semantics: Science, Services and Agents on the World Wide Web* 46 (2017), pp. 2–13 (cit. on pp. 12, 13).

[71]Pek Van Andel and Danièle Bourcier. „Le long voyage d'une notion La sérendipité, de la fiction à la science“. In: *La sérendipité, le hasard heureux. Edited by Danièle Bourcier and Pek van Andel. Paris: Hermann* (2011), pp. 141–59 (cit. on p. 24).

[72]Jörg Waitelonis and Harald Sack. „Towards exploratory video search using linked data“. In: *Multimedia Tools and Applications* 59.2 (2012), pp. 645–672 (cit. on pp. 11, 13, 23).

[73]Jörg Waitelonis, Harald Sack, Johannes Hercher, and Zalan Kramer. „Semantically enabled exploratory video search“. In: *Proceedings of the 3rd international semantic search workshop*. ACM. 2010, p. 8 (cit. on p. 23).

[74]Ryen W White. *Interactions with search systems*. Cambridge University Press, 2016 (cit. on pp. 3, 7, 8, 20–23, 25, 28, 30, 31).

[75]Ryen W White and Resa A Roth. „Exploratory search: Beyond the query-response paradigm“. In: *Synthesis lectures on information concepts, retrieval, and services* 1.1 (2009), pp. 1–98 (cit. on pp. 2–4, 8, 9, 13, 14, 19–23, 26–28, 30, 31, 34).

[76]Ryen W White, Gary Marchionini, and Gheorghe Muresan. „Evaluating exploratory search systems“. In: *Information Processing and Management* 44.2 (2008), p. 433 (cit. on p. 2).

[77]Ryen W White, Steven M Drucker, Gary Marchionini, Marti Hearst, et al. „Exploratory search and HCI: designing and evaluating interfaces to support exploratory search interaction“. In: *CHI'07 extended abstracts on Human factors in computing systems*. ACM. 2007, pp. 2877–2880 (cit. on p. 2).

[78]Ryen W White, Gheorghe Muresan, and Gary Marchionini. „Report on ACM SIGIR 2006 workshop on evaluating exploratory search systems“. In: *Acm Sigir Forum*. Vol. 40. 2. ACM. 2006, pp. 52–60 (cit. on p. 2).

[79]Barbara M Wildemuth and Luanne Freund. „Assigning search tasks designed to elicit exploratory search behaviors“. In: *Proceedings of the Symposium on Human-Computer Interaction and Information Retrieval*. ACM. 2012, p. 4 (cit. on pp. 19–22).

[80]Max L Wilson, MC Schraefel, and Ryen W White. „Evaluating advanced search interfaces using established information-seeking models“. In: *Journal of the American Society for Information Science and Technology* 60.7 (2009), pp. 1407–1422 (cit. on p. 14).

[81]Iris Xie. „Information searching and search models“. In: *Encyclopedia of library and information sciences* (2010), pp. 2592–2604 (cit. on p. 25).