Proceedings of the 55th Hawaii International Conference on System Sciences | 2022

Mapping Knowledge Flows in Exploratory Web Searches

Marcelo Tibau
Federal University of the
State of Rio de Janeiro
marcelo.tibau@uniriotec.br

Sean W. M. Siqueira
Federal University of the
State of Rio de Janeiro
sean@uniriotec.br

Bernardo Pereira Nunes
Australian National University
Bernardo.Nunes@anu.edu.au

Abstract

This paper provides an understanding of the knowledge flow in exploratory Web searches. Based on the Design Science Research epistemology, we represented the information-gathering behaviour and uncovered the knowledge flow of Web searchers as a Knowledge-intensive Process (KiP). By mapping searchers' steps and paths during Web searches, representing their search patterns and decision-making process, it was possible to reveal the knowledge flow and infer the resources more likely to be selected to meet the searchers' information needs. With the help of six teachers, we applied the Think-Aloud method in a scenario where they searched for online (educational) resources on the Web. The application of the Think-Aloud method made knowledge-flow processes explicit throughout Web searches. The results can support new strategies for information retrieval systems and be applied to support expertise exchange and innovation.

1. Introduction

Knowledge flow is an economic concept based on knowledge generation and networks [1]. Its importance to economics, knowledge management and information systems arises from promoting innovation by connecting people to knowledge flows that can be used within and across organisational boundaries. In this sense, information retrieval systems (e.g., Web search engines) open up opportunities to connect knowledge from information networks (e.g., the Web), in which knowledge flows can be transferred both at the individual level (e.g., the searcher) and the community level (e.g., the organisation) [2]. The interaction between the information retrieval system and the searcher is an inner part of the process. A knowledge flow requires an active effort by the searcher to fill gaps in the transmitted knowledge and correct transmission errors [1]. In a way, this interaction can be considered the searching process itself. Web searches provide ways to access knowledge from external channels and apply it within companies' innovation processes to create value [3].

Although any search could be motivated by an information need and a desire for relevant information, not every search involves a complex information problem, nor misunderstandings about terminology and information structure often accompany it. [4]. The knowledge that could be applied within an innovation process usually demands exploratory strategies to enable new associations, discover knowledge, and help searcher's decision-making. It is the essence of a type of search called exploratory search, which is the specialisation of information-seeking that combines activities of querying and browsing to obtain information [4] and, alongside digital technology, emerges as one of the greatest driving forces to provide access to knowledge flows that support innovation [5]. Exploratory searches involve multiple iterations over long periods of time and return results that are thoroughly and critically assessed before being integrated into the searcher's knowledge base [6]. Nevertheless, a lack of skill to design a searching solution or define the problem accurately impacts the search's success because the Web search engine relies on search pattern identification by tracking its interaction with the searcher to narrow the gap between the information it retrieves and its matching to the searcher's information need. Traditionally, search engines provide the search pattern based on keyword matching, which often returns thousands of resources for a single query and makes searching less efficient in some scenarios, such as exploratory searches [7].

New models of search pattern mapping emerged to aid search engines improve their searching efficiency, including the narrow search (helping the user to refine his/her search step by step); the associated search (aiding the user to broaden his/her search by recommending associated contents); the topic-based search (allowing semantic granularity); and, knowledge

URI: https://hdl.handle.net/10125/80007
978-0-9981331-5-7

flow search (mining the association relations among topics) [7].

This paper provides an understanding of knowledge flows in Web exploratory searches. Our approach is focused on mapping searchers' information-gathering behaviour as knowledge-intensive processes (KiP). We argue that modelling a searcher's knowledge flow as a KiP represents a move towards information needs identification. Thus, it is important to clarify the relationship between the concepts of information and knowledge in the context of searching. What is retrieved by a search engine is information (more accurately, pieces of information). The knowledge emerges from the person's searching and interacting with the retrieved results and the cognitive effort applied in this task. Thus, the search process includes both the information retrieval and the knowledge emergence work. For this reason, it is considered that searchers explore knowledge flows instead of information flows [8]. By mapping the steps and paths taken by multiple searchers during a search, we argue it is possible to capture the knowledge flow and infer the resources more likely to be chosen by the searcher to meet her/his information needs.

Having an information retrieval system capable of providing access to knowledge flows (including flows that support innovation) is challenging. Searchers' assessment of particular information differs due to the nature of the work they perform, i.e., the same information is perceived by different searchers differently according to their information needs [9]. A Web search engine can learn from knowledge-intensive searching behaviour, use this knowledge to understand a search intention better, and organise its content accordingly, thus connecting searchers to useful knowledge flows. Searchers would benefit from access to more comprehensive content. As a result, our primary research question is as follow: "Is it possible to map knowledge flows as KiPs?".

The study presented in this paper is grounded on Design Science Research epistemology, which is based on Herbert Simon's distinction between natural sciences, e.g., how things are and how they work, and the sciences of the artificial, e.g., how things should be rearranged to solve specific problems or to allow us to achieve certain goals [10]. We applied the Design Science Research Methodology (DSRM) [11] to study the interactions of artifacts (e.g., search engines) and their context (e.g., exploratory search) and improve our understanding regarding the knowledge flow's entanglement to the searching process. The DSRM has a nominal process sequence of six steps. Table 1 details each of the steps.

This study maps knowledge flows in exploratory search through the Think-Aloud method, in which six teachers from different Italian school levels [12] performed three different exploratory search tasks to find online resources and educational content to be used in their classes. The searching tasks were video-recorded, and the participants verbalised aloud their thoughts, explaining their search approaches and decisions regarding the suitability of the information sources visited.

With videos of the search tasks and the participants' think-aloud verbalisation, under the Design Science Research (DSR) paradigm [10] [13], we represented the knowledge process obtained from the searchers' information-gathering behaviour and the searches' knowledge flows in a Knowledge-intensive Process (KiP) map. As we derive the mapped knowledge flows from exploratory searches, the artifact modelled by the Exploratory Search KiP is named. To the best of our knowledge, this is the first time search processes are deemed knowledge-intensive.

The remainder of this paper is organised as follows. Section 2 presents the theoretical background and related work. Section 3 and 4 presents the planning of the artifact and its development, respectively. Finally, Section 5 discusses the research results regarding the assessment questions and concludes the paper.

2. Theoretical Background

2.1. The Knowledge Flow Theory

The Knowledge Flow theory implies that knowledge moves in an unrestricted manner [8], but to associate knowledge with a purpose , such as innovation, or learning, requires an epistemological perspective[1] that tells apart the knowledge someone has and the actions this person carries out [14]. Having a knowledge purpose is to bridge what is possessed as "knowledge" and what is part of the action of "knowing" [14]. The knowledge flows influence the cognitive activities (e.g., acquisition of knowledge, transfer of knowledge, creation of knowledge) required to enable and shape meaning [15]. Knowledge is increasingly viewed in the frame of both an epistemology[2] of possession and practice [14] [15], then it is necessary to approach any knowledge-related subject as both processes and flows [16].

Conceptually, knowledge flows can be described as dynamic knowledge. They relate to concepts like knowledge conversion, transfer, sharing, integration, and reuse [17]. Rather than focusing on the knowledge

[1]Concerning the nature of knowing and its relationships.
[2]The nature, origin, and scope of knowledge.

Table 1. Steps of the Design Science Research Methodology used in this work (based on Peffers et al., 2007 [11]).

DSRM Step	Description	Reported
(1) Problem identification and motivation	Define the specific research problem and justify the value of a solution.	Section 1
(2) Definition of the objectives for a solution	Infer the objectives of a solution from the problem definition.	Section 3
(3) Design and development	Create the artifact solution.	Sections 3 and 4
(4) Demonstration	Demonstrate the efficacy of the artifact to solve the problem.	Section 4 and 5
(5) Evaluation	Observe and measure how well the artifact supports a solution to the problem.	Sections 3, 4 and 5
(6) Communication	Communicate the problem and its importance, the artifact, its utility and novelty, the rigour of its design, and its effectiveness.	Accounts for the present paper

as a resource, they centre their interest on depicting changes, movements, and applications of knowledge over time [17]. Knowledge flows are associated with their environment (i.e., knowledge space) and the direction they flow [17]. Metaphorically, we can relate knowledge flows to producer and consumer interactions. The producer uses knowledge to create information that is transmitted as data via signals across some knowledge space. In the context of our study, the knowledge space is the Web. Then, the consumer, in our case an online user, interprets the data received and develops information through the incorporation of meaning and context (by searching and interacting with the retrieved results), and finally develops actionable knowledge through some learning mechanism [18, p.69-78].

In the current search engine's blueprint, the inverted index serves as the heavy labourer of modern search engines, encoding term frequencies, term positions, document structure information, and other forms of document metadata [19]. It treats words as uninterpreted tokens, oblivious to morphology, term similarity, and grammar. Recently, inverted indexes have shifted away, favouring dense vector-based indexes, which encode semantically rich document representations [19]. Despite the technical progress, today's search engines are not that different from classical Information Retrieval systems and rely almost entirely on an index-retrieve-then-rank blueprint. Nevertheless, some features help search engines satisfy users' information needs, including coverage (e.g., the number of topics covered), reliability, indexing of whole documents, objectivity (e.g., absence of advertisements or bias), and, filtering options [20]. While some of these features are implementable (e.g., filtering, content indexing), others elude measurement (e.g., coverage and reliability), thus becoming less prone to be adapted to the index-retrieve-then-rank blueprint [19]. This time-tested approach is missing the searching-gathering behaviour that shows what information is added and its corresponding flow. The knowledge flow theory is used as the theoretical foundation of our work because of the relationship between degrees of belief, information added, and the searching process.

The degree of belief depicts how strong a person believes in the truth of a proposition [21, p.1]. Consequently, the higher the degree of belief someone has about something, the higher is his/her confidence about this something's quality of being logically or factually sound. For example, let's say you are searching for the first time about a given topic. The degree of belief you have regarding the search engine's ability to retrieve useful links, the endorsement provided by order of results it produces, and your own assessment about the Website content, provide you with the means to judge the credibility of the information presented.

Regarding the information added, let Γ = { α1, α2,..., αn } be the set of the current full beliefs of a particular user performing an exploratory search. Suppose that what the user believes is consistent (i.e., this person accepts the propositions on her/his set of beliefs as true). The expected gain of information across the search session is maximised when the degree of belief in each possible query formulation or reformulation matches the rescaling of the information it adds to the user's set of full beliefs [22]. In this case, the probability $P(\alpha i, \Gamma)$ is the probability of some information retrieved by the search engine being added to the user's set of beliefs. Therefore, the information added during the search process is probabilistic and explained by the degree of belief. In our view, the knowledge flow theory provides the search pattern frame and the perspective to consider the knowledge gained by searching and interacting with the results as the information added to the set of full beliefs.

2.2. The Knowledge-intensive Process (KiP)

Vaculín et al. [23] defined this particular class of business processes: KiPs are "processes whose conduct and execution are heavily dependent on knowledge workers performing various interconnected knowledge-intensive decision-making tasks. KiPs are genuinely knowledge, information and data-centric and require substantial flexibility at design and run-time". Our research is aligned with most characteristics derived from this definition, such as knowledge-driven, unpredictability, emergence, goal-oriented, event-driven, non-repeatable, constraint and rule-driven.

This study follows the tacit and explicit knowledge framework by Nonaka and Takeuchi [24] regarding the dimensions of knowledge. Tacit knowledge represents the internalised knowledge an individual has – and may or may not be consciously aware of – but uses it to accomplish particular tasks. On the other hand, explicit knowledge represents the knowledge that an individual holds consciously in a form that can easily be communicated to others. KiP helps to convert internalised tacit knowledge into explicit knowledge and make it available for transfer within a group. The knowledge-intensity of a KiP is recognisable by the diversity and uncertainty of process input and output [25]. This definition suggests that process-related knowledge is strongly human-centred and deeply rooted in the creation, co-creation, sharing, transferring, and application of knowledge [26].

2.3. Related Work

Fteimi and Hopf [27] propose an integrative framework for supporting knowledge work processes based on the interaction between humans and Artificial Intelligence. The framework shows a need to investigate how the new forms of knowledge, created and applied by Human-AI interaction, emerge and how they should be managed and integrated into existing working structures. Finally, Fischer and Wunderlich [28] proposed a multi-level framework considering assumptions and elements from traditional productive knowledge work to illustrate influential factors on job performance based on the contingency theory of action. They depicted attributes from three sub-dimensions (individual, job-related, and organisational conditions) and presented a framework for datafication in knowledge-intensive organisations.

The model presented in our study, the Exploratory Search as a KiP, also shares Fteimi and Hopf's and Fischer and Wunderlich's intention to support dynamic work processes (in this case, exploratory searching). However, our work points out the necessity of transcending knowledge workers by focusing on knowledge-intensive processes themselves and understanding the interrelationships between users and search engines. For these reasons, our work was set up to perceive how Information Retrieval systems (especially Web search engines) influence users' propositions, intentions, and activity streams (i.e., series of searching activities the searchers perform). We implemented the conceptual adaptations to KiP described in Section 4.

3. The planning of the artifact

3.1. Definition of the objectives for a solution

As aforementioned, this study considers exploratory Web searches as a Knowledge-intensive Process[3]. It adapts the KiP models to this situation by investigating the knowledge dimension in searching processes, identifying a reference definition of what a KiP represented to Web searching, and analysing which KiP's characteristics changed due to the particular view represented by the proposed artifact. Our solution aims to map the information-gathering behaviour embedded in an exploratory search, the knowledge process prompted by the information gathered and the search's knowledge flows. Here, we build and evaluate a problem-solving artifact designed to understand its environment (Web search engines) and improve it (by mapping knowledge flows and understanding user's Web search patterns). We argue that this knowledge could be employed to adjust retrieval results according to the searcher's Web search pattern, acting as an information-need verifier.

3.2. Design stage

The concepts of Knowledge-Intensive Process Ontology (KIPO) and Knowledge-Intensive Process Notation (KIPN) were used in the design stage. KIPO explores elements from tacit knowledge with a high-level structure composed of five ontologies[4] [29] uses them to represent KiP's features as it already has

[3] According to [26], Knowledge-intensive Processes (KiPs) are an instance of Business Process Management (BPM), Process Management Systems (PMS) and Knowledge Management (KM).

[4] Business Process Ontology (BPO), Collaborative Ontology (CO), Decision Ontology (DO), Business Rules Ontology (BRO) and Knowledge-Intensive ProCess Ontology (KIPCO).

a specific graphic notation to represent tacit knowledge, the KIPN. KIPN is presented and thoroughly explained in [30]. Here, we summarise the main features of a set of diagrams used to represent the search and user perspectives within a KiP.

The first of these diagrams is the KiP Diagram, which indicates constraints to the flow and represents circumstantial events, innovations and decisions. Another diagram is the Socialisation Diagram, which shows how knowledge acquisition and sharing take place within activities. We consider the collaboration-oriented characteristic too restricted for the scope of a Knowledge-intensive Process applied to the exploratory search. We argue that exploratory search may or may not involve collaboration among people. Moreover, it is not restricted to specialists either, since it introduces as process agent the non-specialist or layperson, as it is the case of a searcher that starts a search without previous knowledge about the subject.

There is also the Decision Diagram, which aims to match detailed decision-making processes with their respective results. Finally, there are three last sets of diagrams: (i) Agent Diagram that maps the experience and expertise of agents and illustrates their skills; (ii) Intention Diagram that represents desires, feelings and beliefs that motivate an agent to engage in activities, decisions or socialisation; (iii) Business-Rules Diagram that represents documented business rules that limit a decision during a Knowledge-intensive Process, e.g., legislation, contracts, and internal regulations. Although expertise, intentions, informal regulations and conventions that are part of the professional knowledge of users are very relevant in the context of this work, these sets of diagrams were not used *per se*. Still, we considered this type of tacit knowledge part of the Socialisation and Decision Diagrams instead. Figure 1 shows examples of the Socialisation Diagrams and the Decision Diagram composed with elements from the remaining sets of diagrams.

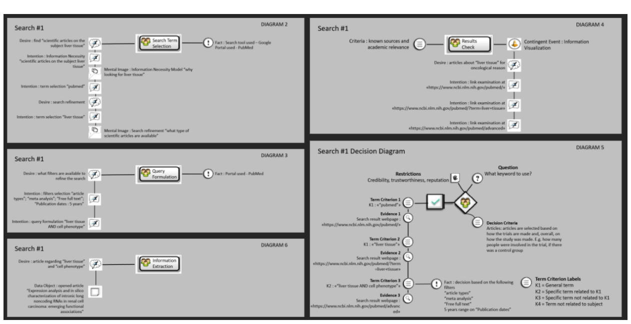

Figure 1. Examples of the mapped diagrams

3.3. Assessment questions

We evaluated the problem-solving artifact based on Hevner's three cycles of activities - Relevance Cycle, Rigour Cycle and Design Cycle [31]. The Relevance Cycle requires that the artifact solves a real-world problem, guiding its field-testing (i.e., the artifact's evaluation guidelines). The Rigour Cycle provides grounding theories, methods, domain experience and expertise, adding the new knowledge derived from the research to a growing knowledge base. Finally, the Design Cycle supports iterative research activities to design, develop and evaluate artifacts and processes.

Following the DSRM systematisation, we intend to evaluate this research by answering three questions: *(i) Did the artifact provide theoretical additions to the literature?*; *(ii) Was the artifact valid?*; and *(iii) Was the problem solved?*. By answering these questions, we apply the three cycles framework [31], focusing on DSR problem-solving paradigm while grounding the effort into scientific knowledge (the theories or scientific literature which grounds the research). The research cycle[5] was carried out by applying the Exploratory Search KiP model in a scenario with expert teachers and validating it with an in-depth analysis of the results and interviews based on the think-aloud protocol[6].

4. The development of the artifact

4.1. KiP's adaptation

To implement a conceptual adaptation to KiP, we first needed to define the knowledge dimension in search processes. As in several knowledge-based human activities, such as reading, driving, learning, and planning, a Web searching process involves information processing. Notably, it involves instructional information, a type of semantic information processing defined for the present purpose as meaningful and well-formed data [32]. It also embraces the concept of need, particularly its instrumentality aspect of reaching the desired goal [33, p.80], as the motivation for information-seeking. Therefore, the knowledge dimension in Web searching requires finding useful information that helps the searcher achieve an information need, exploring and building the knowledge flows.

Next, we had to define what a KiP represented under this particular view of the knowledge dimension

[5] The creation and evaluation of artifacts are an important part of the Design Science Research process, in which artifacts are designed and developed in closely related activities during research cycles.

[6] The protocol requires participants to describe aloud what they are thinking while performing a specific set of tasks.

requirements. Knowledge-intensive Processes are characteristically less rigid in structure and usually involve autonomous user decisions and unpredictable events. Its importance in the Business Process Management domain has emerged due to a prominent role of knowledge workers – a type of worker whose main capital is knowledge, such as engineers, scientists, lawyers and academics – in modern organisations [26]. KiPs are "often related to the need for considering and understanding the knowledge dimension in business processes" [26]. Consequently, the knowledge dimension shifts from the traditional approach that intends to manage processes and process-related knowledge separately towards an approach that considers it an integral part or the process itself. Hence, the diagrams show the knowledge flow created by the collaboration between the agents. Finally, we analysed which KiP's characteristics changed due to our particular view of KiP and its knowledge domain's requirement. In this sense, while applying KiP to exploratory search, we noticed one more aspect not considered by its KM inheritance: knowledge acquisition.

4.2. Exploratory Search KiP model - first iteration

To expose the Knowledge-intensive activities and the searcher's characteristics, we defined a few questions to guide the mapping effort of behaviours and tasks: *How were the initial search terms chosen? How were the search terms evaluated? How and why were search terms modified during a search session? How were the results checked? What was the Decision Process involved? What features determined data suitability?*

As an initial attempt, searches were conducted based on Kules and Shneiderman [34]. In their design guidelines for categorised overviews, the authors proposed a scenario and a search task to help identify search patterns. Kules and Shneiderman's scenario was originally designed to stimulate the planning of a search activity. After, we used its assignment to produce a list of ideas for a series of articles on the aging workforce as a search task. This first iteration was performed before the think-aloud study. The result of this earlier work was the definition of four exploratory search activities, as shown in Figure 2: (i) Search Term Selection, (ii) Query Formulation, (iii) Results Check[7], and (iv) Information Extraction.

[7]The activity was named "Results Exam" in the first version shown in Figure 2. It changed in later versions to "Results Check". We opted to present the activity with its current name in the text.

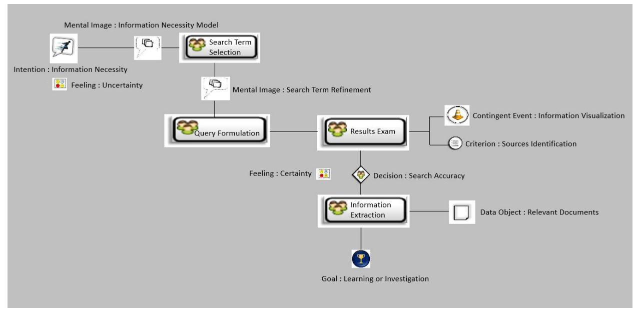

Figure 2. Exploratory Search KiP Model version one

4.3. The Think-Aloud Method

Think-aloud is a method in which participants speak aloud any words that come into their minds as they complete a task [35]. As a traditional method in qualitative research, the think-aloud protocol is a concurrent verbalisation procedure [36], in which the information is verbalised at the time the participant is attending to it. As a method, it is rooted in cognitive psychology and considers that verbal reports are data. It is applied in situations where verbalising is subordinated to (or dependent on) the cognitive process [36]. However, based on the concept of *inner speech* [37], verbalising protocols are usually not expressed in complete and reasoned sentences since the subject of the *talk* is usually evident to the person speaking. The participants are then asked to report verbal or non-verbal content [38], known as type 1 and type 2 verbalisation, but usually do not need to explain their cognition (type 3 verbalisation). At the protocol used in this research, the teachers were asked to verbalise their thoughts, feelings, and cognitive processes while performing their searches, explaining each decision made during their exploratory search tasks to include all three types of verbalisation.

According to previous researches [39, p.156], five participants are sufficient for an effective usability test using the think-aloud protocol. However, Lewis and Mack [40] had six participants to show the evidence of thinking aloud protocols in a learning situation on an information system. We, therefore, decided to perform this study with six teachers too.

4.4. Exploratory Search KiP model - second iteration

The second iteration used data from a think-aloud study, in which six teachers from different Italian school levels were selected [12] and observed while searching online resources and educational content to be used in their classes. Videos showing the teachers' interactions with search engines and the think-aloud protocol served as a basis for analysing the model's effectiveness in representing the knowledge processes and flows presented by these search interactions.

The teachers were given three searching tasks. The first one was to open a portal they often use to find teaching resources and repeat a search they had previously performed and considered a successful one, i.e., a retrieved teaching resource used in the classroom. The goal was to understand what was considered a successful search and the steps taken to achieve it. As for the second task, they were asked to return to the same portal to look for a new resource or topic they had not investigated yet, but would be interested in. The goal of this particular task was to verify the occurrence of a Web search pattern that could be responsible for the search's success. Finally, the third task was to look for a completely new resource using the teacher's Web search engine of choice. This task aimed to observe the entire search process and how the users applied the steps, and the patterns used in both previous tasks. The transcription of the searches from the think-aloud protocol allowed a detailed analysis regarding the reasons for the teachers' behaviours during the searches, provided insights into their motivation while performing the activities, and aided to map the search movements operationally (e.g., the sequences in which the teachers performed their actions).

To provide the reader with a better understanding of the modelling process, we briefly describe how the searcher's information need was identified (at the model, the term was named "desire" and depicted the goal that motivated the search). We also provide additional clarification about the process in which the "desire" was transformed into action. It was characterised in the model by the term "intention". Moreover, we describe how each action was used to determine the "mental image" that represented how the searcher's awareness about the subject evolved. The teacher stated her/his "desire" before starting the search, e.g., "I want to find scientific articles". When the teacher decided to open a Web search engine or a portal and chose a search term to begin her/his search, it was considered that s/he had passed from the desire to intention - e.g., the usage of the term "liver tissues" to formulate a query. The psychological change of state from desire to intention occurred when the teacher decided to click on a certain link during the results check. To consider that a mental image was formed (or refined), three criteria were used. The first one, a subjective criterion, was used when the results check created a questioning declared by the teacher, e.g., "which articles are available on this topic?". The second was applied when the user decided to refine or reformulate the queries, e.g., from "coaching with compassion" to "coaching with compassion scale OR questionnaire OR questionnaires OR questions". The third, an objective criterion, was used when the teacher clearly stated her/his awareness of a situation or fact prompted by the search, e.g., "I understood this topic" or "this will help me". Generally, the third criterion situation was reinforced by information extraction, e.g., a file downloaded or a link marked as favourite.

Initial mappings encountered difficulties capturing the relevance of context variables, i.e., what type of resource the user was looking for and for what purpose, and the teachers' expertise used during the searching activities as search filters. These difficulties were mitigated by review processes that were conducted to get the searchers' feedback. Every mapping was thoroughly analysed to identify possible divergences between what the teacher said and what was done during the search. The persons involved were then asked to provide more information about the particular search. The same happened in situations where no clear evidence about a mental image representation was perceived. This process aided in refining the Diagrams. Figure 3 shows an example of such interactions.

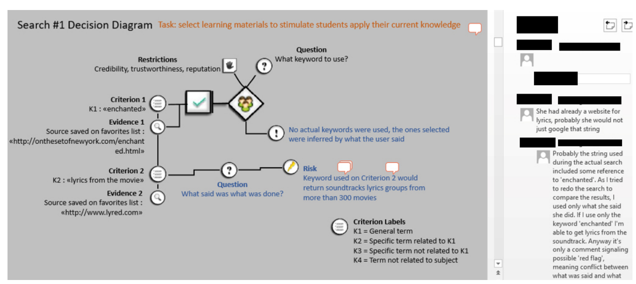

Figure 3. Example of a review process conducted on one of the mapped searches

During the think-aloud protocols, the teachers' explanations led to the inclusion of labels and comments to better capture, and represent their actions and words. It was possible to see that the decisions involved in the activities were, in a general sense, repeated. They were used to stimulate terms selection and define the results examined, generating two decision criteria adopted by the teachers throughout the searches.

The first decision criterion, concerning keywords, was named "Term Criterion Labels", and the labels were categorised into four kinds[8]. The second type involved criterion related to the decision about what link to click and what resource to choose and was named "Decision Criteria". It is noteworthy that the

[8] K1 (General term); K2 (Specific term related to K1); K3 (Specific term not related to K1); and K4 (Term not related to the subject - representing a new search).

two criteria types would not be detectable without the visualisation provided by the Exploratory Search KiP model. These criteria mirrored the teachers' expertise in their knowledge domain and experience about their students' personalities and class context. Figure 4 shows the second version of the artifact with its four activities explained.

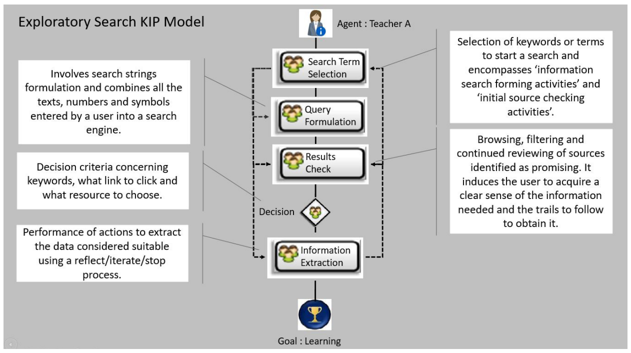

Figure 4. Exploratory Search KiP Model - second version

5. Discussion and Conclusion

The assessment questions proposed in the subsection 3.3 are clues to validate whether theoretical additions were made (Rigour Cycle), the artifact was valid (Design Cycle), and the problem was solved (Relevance Cycle). The Exploratory Search KiP model applied a BPM approach to a non-business situation (Web searches). By doing so, KiP's business aspects - such as being heavily dependent on knowledge workers, Person-to-Person (P2P) type of interaction, and collaborative multi-user environment - were adapted. When applied to Web search activities, the model represented the search process by its knowledge flow. By expanding KiP's ability to represent knowledge and enabling a Person-to-Machine (P2M) interaction, adapting KiP's literature, the model provided a positive answer to the first assessment question ("Did the artifact provide theoretical additions to the literature?").

The Knowledge-intensive activities from the Exploratory Search KiP model ("Search Term Selection", "Query Formulation", "Results Check", and "Information Extraction") helped to provide more visibility to the behaviours involved in an exploratory search process. By visually describing the relationships between knowledge acquisition, sharing, storing and reusing, the model made it possible to distinguish three main behaviours involved in the mentioned activities: **(i) The first behaviour**, subject familiarisation, is observable when the searcher becomes familiar with the subject. It helps to reduce the uncertainty involved in the activity and makes it possible for the searcher to select useful search terms (as seen in Fig. 3 with the criterion and evidence labels); **(ii) The second behaviour** involves the ability to control the search process itself, and it increases as it moves from the first KiP activity to the second. Queries formulation and reformulation, formalised by search string definition, indicates an increased awareness about the topic (e.g., Fig. 2 captured the uncertainty/certainty movement with the icon "feeling"). It is impossible to define a string if there is no conceptual understanding whatsoever about the searching subject, and **(iii) The third behaviour** encompasses the rest of the Exploratory Search KiP model and is characterised by an ability to assess the retrieved information relevance. This evaluation capability influences what information the searcher selects, extracts and stores (e.g., the mapped diagrams shown in Fig. 1). The Exploratory Search KiP model positively answered the second assessment question ("Was the artifact valid?").

The answer to the third assessment question ("Was the problem solved?") is not straightforward. Web search engines' information retrieval is based on link analysis that assigns a weighting to each element in a set of documents. The weight provides importance within the set with criteria that could be either (a) by estimating the number and the quality of the links referring to the element; or (b) by semantic proximity of the searching concepts. Either way, it seeks to match keywords and terms to documents (using individual information such as bookmarks and search history as set-up features). Still, it does not consider searchers' information-gathering patterns during search activities. By providing a map showing the information-gathering behaviour during search activities (e.g., Figures 1, 2, 3, and 4), the Exploratory Search KiP model positively answered the third assessment question.

Although the three assessment questions were satisfactorily answered, the lack of an implemented Information Retrieval ranking solution to determine document relevancy based on the searching behaviours limits this research.

In conclusion, we would like to answer our primary research question: Is it possible to map knowledge flows as KiPs? Indeed the Exploratory Search KiP model was able to map the knowledge flows from the teachers' searches. As it enabled the present study to visualise the teachers' search pattern and map their decision-making process, the Exploratory Search KiP model might be used to understand the Web search information need and knowledge flow processes, as commented in Section 1. The visualisation also could be applied to support expertise exchange and be used in studies in which the expert domain has a pivotal role in the process. The Exploratory Search KiP model could also support developing a new generation of intelligent tutoring systems grounded on understanding learners' decision-making while interacting with the system. We envision developing a ranking solution based on the searchers' behaviours and information needs as future work. Having an information retrieval tool with this approach would allow us to apply it to real-world business scenarios to analyse the knowledge flows and the innovation process based on the knowledge gained by in-company Web searches and how it is used to create value.

Acknowledgement

This study was financed in part by the "National Council for Scientific and Technological Development (CNPq) - Brazil" - Process 315374/2018-7, Project "Searching as Learning: the information search as a tool for learning" and by the "Coordination for the Improvement of Higher Education Personnel" (CAPES) – Brazil – Finance Code 001.

References

[1] O. Sorenson, J. W. Rivkin, and L. Fleming, "Complexity, networks and knowledge flow.," *Academy of Management Proceedings*, vol. 2004, no. 1, pp. R1–R6, 2004.

[2] D. Tessier, "Enabling knowledge flow: The knowledge management triangle model," in *Handbook of Research on Organizational Culture Strategies for Effective Knowledge Management and Performance* (D. Tessier and M. Jennex, eds.), ch. 9, pp. 157–175, IGI Global, 2021.

[3] Y.-H. Chiang and K.-P. Hung, "Exploring open search strategies and perceived innovation performance from the perspective of inter-organizational knowledge flows," *R&D Management*, vol. 40, no. 3, pp. 292–299, 2010.

[4] R. White and R. Roth, "Exploratory search: beyond the query-response paradigm," Synthesis lectures on information concepts, retrieval, and services, 1(1): 1-98. Print., 2009.

[5] Y. Liao, E. R. Loures, F. Deschamps, G. Brezinski, and A. Venâncio, "The impact of the fourth industrial revolution: a cross-country/region comparison," *Production*, vol. 28, 00 2018.

[6] G. Marchionini, "Foundations for personal information infrastructures: information-seeking knowledge, skills, and attitudes," in *in Information Seeking in Electronic Environments (Ch.4)*, pp. 61–75, New York NY: Cambridge University Press, 1995.

[7] X. Wei, X. Luo, Q. Li, and J. Zhang, "Knowle: Searching news in the search pattern of knowledge flow," in *Web Information Systems Engineering – WISE 2013* (X. Lin, Y. Manolopoulos, D. Srivastava, and G. Huang, eds.), (Berlin, Heidelberg), pp. 527–530, Springer Berlin Heidelberg, 2013.

[8] D. Wang, "Knowledge flow as a metaphor? a grounded theory of information and knowledge networks," *Academy of Management Proceedings*, vol. 2012, no. 1, p. 16722, 2012.

[9] F. J. Devadason and P. P. Lingam, "A methodology for the identification of information needs of users," *IFLA Journal*, vol. 23, no. 1, pp. 41–51, 1997.

[10] H. A. Simon, *The Sciences of the Artificial (3rd Ed.)*. Cambridge, MA, USA: MIT Press, 1996.

[11] K. Peffers, T. Tuunanen, M. Rothenberger, and S. Chatterjee, "A design science research methodology for information systems research," *J. Manage. Inf. Syst.*, vol. 24, p. 45–77, Dec. 2007.

[12] M. Bortoluzzi and I. Marenzi, "Web searches for learning: How language teachers search for online resources," *Lingue e Linguaggi*, vol. 23, pp. 21–36, 12 2017.

[13] S. Gregor and A. R. Hevner, "Positioning and presenting design science research for maximum impact," *MIS Q.*, vol. 37, pp. 337–356, June 2013.

[14] S. D. N. Cook and J. S. Brown, "Bridging epistemologies: The generative dance between organizational knowledge and organizational knowing," *Organization Science*, vol. 10, no. 4, pp. 381–400, 1999.

[15] A. Aribi and O. Dupouët, "Absorptive capacity: a non-linear process," *Knowledge Management Research & Practice*, vol. 14, no. 1, pp. 15–26, 2016.

[16] D. Chauvel, "Knowledge as both flows and processes. Proposed by GeCSO 2013 conference committee," *Knowledge Management Research & Practice*, vol. 14, pp. 1–3, Feb. 2016.

[17] M. E. Nissen, "Knowledge management and global cultures: elucidation through an institutional knowledge-flow perspective,"

Knowledge and Process Management, vol. 14, no. 3, pp. 211–225, 2007.

[18] G. Von Krogh, K. Ichijo, I. Nonaka, *et al.*, *Enabling knowledge creation: How to unlock the mystery of tacit knowledge and release the power of innovation*. Oxford University Press, 2000.

[19] D. Metzler, Y. Tay, D. Bahri, and M. Najork, "Rethinking search: Making domain experts out of dilettantes," *SIGIR Forum*, vol. 55, July 2021.

[20] J. Liu, H. Zamir, Y. Li, and S. K. Hastings, "Search systems and their features: What college students use to find and save information," *Library Information Science Research*, vol. 40, no. 2, pp. 118 – 124, 2018.

[21] F. Huber, "Belief and Degrees of Belief," in *Degrees of Belief* (F. Huber and C. Schmidt-Petri, eds.), pp. 1–33, Dordrecht: Springer Netherlands, 2009.

[22] P. Milne, "Probability as a Measure of Information Added," *Journal of Logic, Language and Information*, vol. 21, pp. 163–188, Apr. 2012.

[23] R. Vaculín, R. Hull, F. Heath, C. Cochran, A. Nigam, and N. Sukaviriya, "Declarative business artifact centric modeling of decision and knowledge intensive business processes," pp. 151 – 160, 10 2011.

[24] I. Nonaka and H. Takeuchi, *The knowledge-creating company: how Japanese companies create the dynamics of innovation*. New York: Oxford University Press, 1995.

[25] T. H. Davenport, S. L. Jarvenpaa, and M. C. Beers, "Improving Knowledge Work Processes," *Sloan Management Review*, vol. 37, no. 4, pp. 53–65, 1996.

[26] C. Di Ciccio, A. Marrella, and A. Russo, "Knowledge-intensive processes: Characteristics, requirements and analysis of contemporary approaches," *Journal on Data Semantics*, vol. 4, pp. 29–57, 03 2015.

[27] N. Fteimi and K. Hopf, "Knowledge Management in the Era of Artificial Intelligence - Developing an Integrative Framework," *ECIS 2021 Research-in-Progress Papers*, June 2021.

[28] L. Harder Fischer and N. Wunderlich, "Datafication of knowledge work: A framework and research agenda," in *54th Hawaii International Conference on System Sciences (HICSS 2021)*, (Hawaii, USA), 01 2021.

[29] J. França, J. Netto, J. Carvalho, F. Santoro, F. Baião, and M. Pimentel, "Kipo: the knowledge-intensive process ontology," *Software and Systems Modeling*, vol. 14, pp. 1127–1157, 07 2015.

[30] J. Netto, F. Santoro, and F. Baião, "Evaluating kipn for modeling kip," pp. 549–561, 05 2014.

[31] A. Hevner, "A three cycle view of design science research," *Scandinavian Journal of Information Systems*, vol. 19, 01 2007.

[32] N. Fresco and M. J. Wolf, "The instructional information processing account of digital computation," *Synthese*, vol. 191, no. 7, pp. 1469–1492, 2014.

[33] D. O. Case and L. M. Given, *Looking for information: a survey of research on information seeking, needs, and behavior*. Studies in information, Bingley, UK: Emerald, fourth edition ed., 2016. OCLC: ocn923550250.

[34] B. Kules and B. Shneiderman, "Users can change their web search tactics: Design guidelines for categorized overviews," *Inf. Process. Manage.*, vol. 44, pp. 463–484, Mar. 2008.

[35] E. Charters, "The use of think-aloud methods in qualitative research an introduction to think-aloud methods," *Brock Education : a Journal of Educational Research and Practice*, vol. 12, pp. 68–82, 05 2010.

[36] K. Ericsson and H. Simon, "Verbal reports as data.," *Psychological Review*, vol. 87, no. 3, pp. 215–251, 1980.

[37] L. S. Vygotsky, *Studies in communication. Thought and language, (E. Hanfmann G. Vakar, Eds.).* Cambridge, MA, USA: MIT Press, 1962.

[38] S. Vandevelde, H. V. Keer, G. Schellings, and B. V. Hout-Wolters, "Using think-aloud protocol analysis to gain in-depth insights into upper primary school children's self-regulated learning," *Learning and Individual Differences*, vol. 43, pp. 11 – 30, 2015.

[39] J. Nielsen, *Usability Engineering*. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 1994.

[40] C. Lewis and R. Mack, "Learning to use a text processing system: Evidence from "thinking aloud" protocols," in *Proceedings of the 1982 Conference on Human Factors in Computing Systems*, CHI '82, (New York, NY, USA), p. 387–392, Association for Computing Machinery, 1982.