THE PHILOSOPHICAL APPROACH TO INFORMATION SEEKING AND RETRIEVAL CALLED HERMENEUS

Fabiano Duarte Beppler
Ph.D. in Knowledge Engineering and Management from The Federal University of Santa Catarina, Brazil. Founding Partner of KNEWIN - Intelligence in Information Retrieval, Brazil, and Researcher of Stela Institute, Brazil.
E-mail: fbeppler@stela.org.br

Frederico Torres Fonseca
Ph.D. in Spatial Information Science and Engineering from The University of Maine, USA. Associate Professor of The College of Information Sciences and Technology at the Pennsylvania State University, USA.
Email: fredfonseca@ist.psu.edu

Roberto Carlos dos Santos Pacheco
Ph.D. in Production Engineering from The Federal University of Santa Catarina, Brazil. Adjunct Professor of The Federal University of Santa Catarina, Brazil, and Lead Researcher of Stela Institute, Brazil.
E-mail: pacheco@stela.org.br

Abstract
Information seekers always have a question in mind even when they do not know exactly what it is. How an information seeking and retrieval system can support users while they are still shaping their information needs? We resorted to the hermeneutic circle philosophical concept of Heidegger and Gadamer to provide the principles of such a system. In our implementation users are able to develop their ideas while browsing the information and the concepts that represent the information. We chose ontologies to implement this hermeneutic approach. Ontologies improve information retrieval systems regarding its retrieval and presentation of information, which make the task of finding information more effective, efficient, and interactive. We called our framework Hermeneus, which in Greek means the interpreter or translator. Hermeneus works as an intermediary that facilitates the user to move from the initial state of information need to the goal state of resolution. Our system intends to be the bridge between the user's question and the answer to be found while she or he navigates in the ontology concepts and the instances of these concepts in a back and forth way.

Keywords: Information seeking. Ontology. Hermeneutic circle. Hermeneus.

1 INTRODUCTION

This paper is dedicated to explore and reflect on a philosophical approach to information seeking and retrieval proposed by the authors, with first results published in the *DataGramaZero* - Journal of Information Science (BEPPLER; FONSECA; PACHECO, 2009), based PhD thesis in Knowledge Engineering and Management developed by Beppler (2008).

The information seeker always has a question in mind even when he or she does not know exactly what it is. "How is one to give a name to what he is still searching for? To assign the naming word is, after all, what constitutes finding" (HEIDEGGER, 1982 p. 20). What people do when they are looking for information? How do people understand? How do people change their minds while engaged in information-seeking behavior? How do people know that they found what they were looking for? What compels them to look for more? Finally, how can we address these points in the implementation of an information retrieval system?

The inability to express precisely what information is needed induces users to engage in an indeterminate sequence of interactions with an information system (MARCHIONINI, 1989; KUHLTHAU, 1993; SONNENWALD, WILDEMUTH et al., 2001; CAPRA III; PÉREZ-QUIÑONES, 2005; JANSEN, BOOTH et al., 2008). This challenging position is "an anomalous state of knowledge, which prompts the person to engage in active information-seeking behavior" (BELKIN; CROFT, 1992 p.2). The information-seeking process "includes recognizing and interpreting the information problem, establishing a plan of search, conducting the search, evaluating the results, and if necessary, iterating through the process again" (MARCHIONINI, 1989 p.54). Furthermore, the cognitive model of users may be "dynamic but not self-contained"(INGWERSEN; JÄRVELIN, 2005a p.9). The user's state of knowledge changes by virtue of engagement with retrieved information that by its turn reflects in some change in the anomalous state of knowledge (BELKIN, 1993).

Human behavioral studies show that the information seekers, when using information systems, are lazy to compose queries and rarely adopt more complex expressions in their criteria (SPINK; WOLFRAM et al., 2001). A large percentage of users do not go beyond the initial results, i.e., they have low tolerance to go deep through what was retrieved (SILVERSTEIN; HENZINGER et al., 1999; JANSEN; BOOTH et al., 2008). Furthermore, in order to develop an information-seeking behavior, users have to intervene and interact with retrieved documents and with the representations of documents and their relationships (BELKIN, 1993; INGWERSEN; LARSEN et al., 2010). Therefore, this development of an information-seeking behavior requires systems that should provide high interactivity, besides being more intuitive and responsive (JÄRVELIN; INGWERSEN, 2004; INGWERSEN; JÄRVELIN, 2005b).

In trying to find an answer for some of these questions in an information seeking context, we resorted to Philosophy. More specifically, we used Heidegger's (1996) and Gadamer's (1989) hermeneutic circle as the main principle behind our information retrieval system. The hermeneutic circle is the concept used by Heidegger and Gadamer to explain how humans understand. Before approaching any subject or question, we need to have an idea (the fore-structure) of what the thing is beforehand. In the hermeneutic circle we go back and forth between the thing itself and our conceptualization of it till we finally grasp its meaning.

Our approach is designed to help the user through the information seeking process. The system is flexible enough to allow for changes in the user's ideas and concepts during the process. So, instead of investing in query composition, we invested in developing a structure that lets the user build his or her own query. Furthermore, this query building is not static but dynamic following the hermeneutic circle principles. In our approach the user can see partial results, go back and forth between concepts and instances of these concepts, and change modes of visualization based either on concepts or instances of these concepts.

The architectural option we chose to implement this hermeneutic approach to information seeking and retrieval was to use ontologies to represent the concepts and their relationships. Ontologies improve information retrieval (IR) systems regarding its retrieval and presentation of information. Ontologies can enhance the search process making the task of finding information more effective, efficient, and interactive. In this context ontologies are considered an engineering artifact, constituted by a specific vocabulary used to describe a certain reality, plus a set of explicit assumptions regarding the intended meaning of the vocabulary words (GUARINO, 1998). As an engineering artifact, ontologies may be used to generate or validate information systems components (FONSECA, 2007).

We called our framework Hermeneus following Kuhlthau (1991, 1993)who considered that the IR system is one *intermediary* that should facilitate the user to move from the initial state of information need to the goal state of resolution. Hermeneus in Greek means the interpreter or translator (KINGSLEY, 1993). Our system is intended to be the bridge between the user's question and the answer that she or he will find in the back and forth movement between the concepts and the instances.

Next, in order to show how users of IR systems can come to new insights about the information they are browsing, we describe the hermeneutic circle applied to the information seeking process. In section 3, we present and demonstrate our framework contextualizing its concepts in a prototype and also describe the role of ontologies in our approach. We discuss Hermeneus in section 4, presenting the importance of hermeneutic circle and ontology, besides detailing some points about interaction with additional information. In section 5, related works are presented and, finally, in section 6, we report our conclusions and an outlook to future work.

2 PHILOSOPHICAL FOUNDATIONS: THE METAPHOR OF THE HERMENEUTIC CIRCLE APPLIED TO THE INFORMATION SEEKING PROCESS

The original meaning and scope of hermeneutics as the art of interpretation of texts was greatly extended by the work of Heidegger and Gadamer (GADAMER, 1989; HEIDEGGER; STAMBAUGH, 1996). Heidegger explains that "the expression 'hermeneutic' derives from the Greek verb *hermeneuein*. That verb is related to the noun of *hermeneus*, which is referable to the name of the god Hermes by a playful thinking that is more compelling than the rigor of science. Hermes is the divine messenger. He brings the message of destiny; *hermeneuein* is that exposition which brings tidings because it can listen to a message" (HEIDEGGER, 1982, p. 29).

We think that hermeneutic play is the context in which the user of an IR system will come to new insights about the information she or he is browsing. To be effective, designers of IR systems must structure a context that allows users to engage in the spontaneity of interpretive play. In our framework for information seeking and retrieval we took into account the four main components of hermeneutics and used each one of them to support the components in our framework, as can be seen in Table 1.

Table 1 - Components of hermeneutics related to components of the system

Components of hermeneutics	**Components in the system**
Fore-structure	Computational ontology
Thing in itself	Instances
Play	User interaction with the instances and with the computational ontology
Application	The user's question

First, there is the *fore-structure* (which Gadamer calls prejudices), which is the initial idea that the user has about the subject. This is the first instrument that a user has to start browsing a data set. In our framework fore-structure corresponds to computational ontologies. Second, there is the *thing-in-itself* or things themselves. This corresponds to the objective world, the facts we have; the data we have gathered. In our framework, these are the instances of the concepts in the ontology. Finally, there are *play* and *application*. Here we consider that the user interaction with the instances and with the concepts in the ontology is what represents *play*. Users go back and forth between the two (instances and concepts) always having an *application* in mind.

Heidegger says that the process of understanding involves a continual back and forth dialogue between the dimensions of analysis and synthesis aimed at apprehending the whole, and the dimensions that aim at apprehending the parts. We can never avoid the fact that one comes to the project of understanding with assumptions about the whole, and/or the parts.

But one can, and must, reflect on those presuppositions, while keeping the object of inquiry, "the thing itself" clearly in view. Heidegger says that our "first, constant, and last task is not to let fore-having, fore-sight, and fore-conception be given to it by chance ideas and popular conceptions, but to guarantee the scientific theme by developing these in terms of the things themselves" (1996, p. 143). The process of understanding something new is "a continuous dialectical tacking between the most local of local detail and the most global of global structure in such a way as to bring both into view simultaneously" (GEERTZ, 1979) cited in (Bernstein, 1983).

Heidegger says that "every interpretation which is to contribute some understanding must already have understood what is to be interpreted" (1996, p. 142). Heidegger understands that this proposition may be seen as a vicious circle. "But if interpretation always already has to operate within what is understood and nurture itself from this, how should it then produce scientific results without going in a circle, especially when the presupposed understanding still operates in the common knowledge of human being and world?" (HEIDEGGER; STAMBAUGH, 1996 p. 143) But Heidegger firmly denies that this is a vicious circle. He says that "the fulfillment of the fundamental conditions of possible interpretation rather lies in not mistaking interpretation beforehand with regard to the essential conditions of its being done" (HEIDEGGER; STAMBAUGH, 1996 p. 143). Furthermore, "this circle of understanding is not a circle in which any random kind of knowledge operates" but it is instead the expression of our most fundamental and intuitive knowledge. So following Heidegger's suggestion of "what is decisive is not to get out of the circle, but to get in it in the right way;" we built Hermeneus as one of the ways of getting into the hermeneutic circle and enabling the information seeker to achieve what she or he was looking for.

To Heidegger's description of the hermeneutic circle, Gadamer (1989) has added two crucial elements: *application* and *play*. Instead of conceiving the process of understanding in terms of analysis and synthesis alone, he has shown the importance of the dimension of application to that process. Understanding takes place in a context in which the concepts involved enter into some sort of practice. This requires a judgment about the implications for alternatives possible in local situations. Gadamer has also pointed to the inevitably playful, to and fro, nature of the hermeneutic process. In play, the participant is no longer trapped in a subjectivism in which the play is a kind of predicate over against the participating subject. Instead, the player loses him or herself in the play, becoming, as it were, the predicate of the play. The player's movements and decisions are constrained to conform to the regenerative, to and fro patterns of the play. Our framework for information seeking and retrieval tries to give users the alternatives mentioned by Gadamer.

One of the foundations of Gadamer's work in hermeneutics is the connection between the theoretical and the practical. Hermeneutic understanding is not a selfless, purposeless activity. For Gadamer, interpretation, understanding, and application are all linked together in a unity. These are not independent events in time but they happen simultaneously, with one influencing and actually enabling the other. Gadamer starts saying that "interpretation is not an occasional, post facto supplement to understanding; rather, understanding is always interpretation, and hence interpretation is an explicit form of understanding" (GADAMER, 1989). Understanding and interpretation are applied to a text in Gadamer's example, and applied here to a specific subject domain, which is the target of the user's initial question. The task of the interpreter is to adapt the text to a concrete situation to which the text is speaking. The same is valid here in the process of understanding a subject through the process of going back and forth between the concepts and instances. Understanding is a dynamic process and meaning is always being created and recreated. The objective of our framework is to give the user the medium to achieve this kind of understanding. In the development of our information seeking and retrieval interface we applied the metaphor of the hermeneutic circle to represent the user activity of going back and forth between instances and concepts while he or she is trying to shape his or her own questions.

In the next section we start making the shift from the philosophical foundations of our framework to its computational aspects. We start with a general description of Hermeneus. Then we deal with the main element in the translation of the hermeneutic circle to an IR interface, the ontology. In our approach the ontology corresponds to the fore-structures and has the role of anchoring the information seeking process.

3 A FRAMEWORK TO INFORMATION SEEKING AND RETRIEVAL

Information search process is the user's constructive activity of finding meaning from information in order to extend his or her state of knowledge on a particular problem or topic (KUHLTHAU, 1991). This activity is consistently interactive because users are constantly expanding their state of knowledge along with the search process. Based on such issues, it is clear that IR systems should provide mechanisms to facilitate both the users' interaction and the state of knowledge expansion.

3.1 Hermeneus

We took into account the above issues and developed a framework to information seeking and retrieval called Hermeneus that uses ontologies to enhance users' interactivity with IR systems. Our framework is composed of three interactive components: *ontology navigator*, *retrieved instance*s, and *additional information*. In the *ontology navigator component* (#1 in **Figure 1**), where the ontology is graphically available, users can interact with ontology concepts and visualize how they are related to each other. In addition, users can select a *central concept* in the ontology to designate which kinds of instances are being retrieved (e.g., if *central concept* is *paper*, then the main retrieved instances are papers). In the prototype used to validate our approach we created a database using 622 papers from the Journal of the American Society for Information Science and Technology (JASIST). Each instance contains information about the papers including title, authors, institutions, issues, and venue. Based on the content of such citations, it was conceived an ontology composed of five classes with relations among them. The classes and their respective properties represent the concepts of *paper, author, institution, journal,* and *keyword.*

Users can interact with all retrieved content available in the *retrieved instances* (#2 in **Figure 1**) and *additional information* (#3 in **Figure 1**) components. In the *retrieved instancescomponent*, users can visualize instances of a central concept and instances of directly related concepts. So, for example, in an ontology with three concepts, *paper, author*,and *date*, where there is a relation between the *paper* and the *date* concepts but not between the *author* and the *date* concepts, and the *central concept* is *date*, only instances about *date* and *paper* are presented. In addition, when users select an instance a new search is performed considering such content, like a refinement process. In the *additional information component*, this function may work differently, because such a component can be configuredso as to define specific behaviors for extracted information. For example, in our prototype, if we click in the amount of papers (i.e., eighteen), available in the *additional information component*, related to *Frederico Fonseca*'s paper, we can automatically visualize eighteen papers in the *retrieved instance component*, published by authors linked to the *Pennsylvania State University* (i.e., the affiliation of *Frederico Fonseca*author).

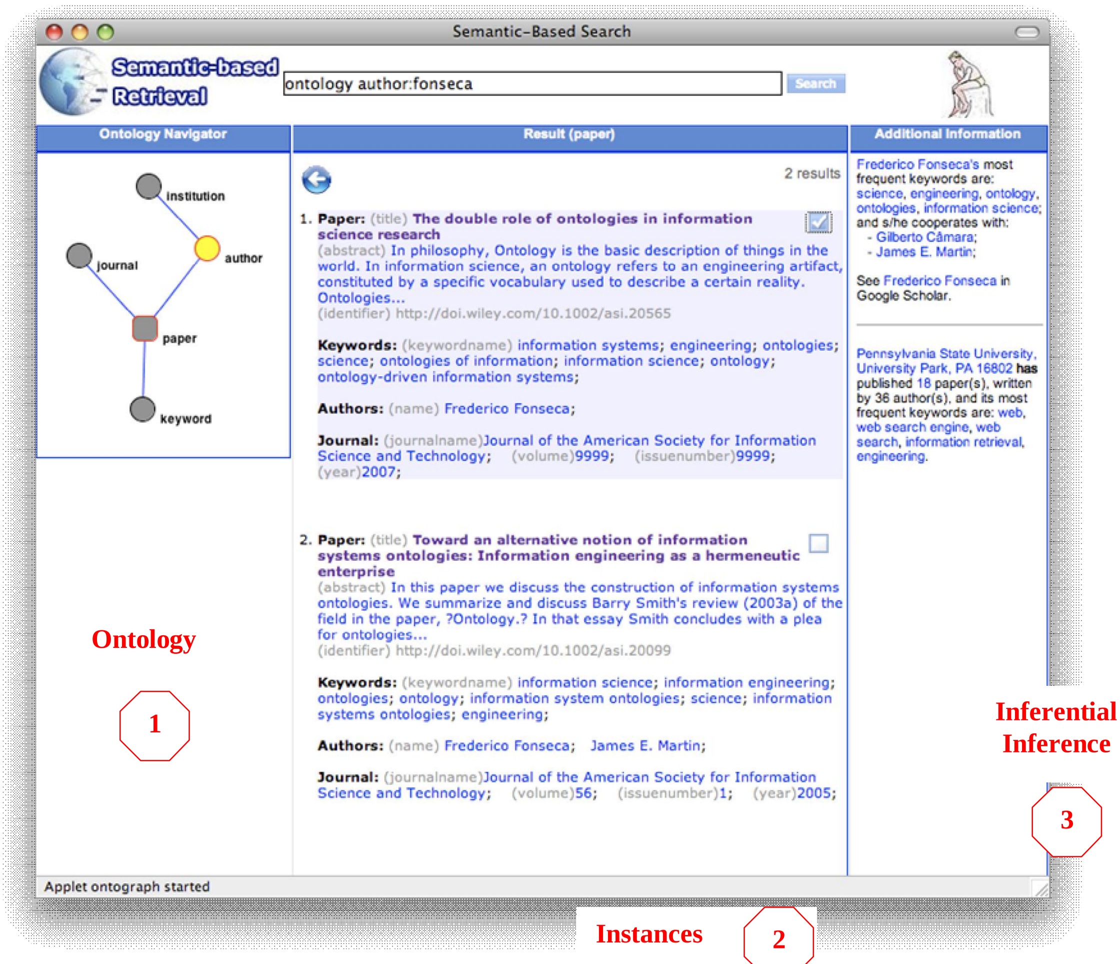

Figure 1 - Prototype's result screen with *ontology navigator* (1), *retrieved instances* (2), and *additional information* (3) components

3.2 The role of ontology in the architecture

Ontology with standard IR tools can be used to create enhanced search capabilities (McGUINNESS, 1999). Ontologies can also be used to provide more accurate answers to queries because with ontologies knowledge is formally represented and users are able to compose contextualized queries (DECKER et al., 1999; MELGAR et al., 2010). Furthermore, as ontologies include axioms to express relationships between concepts, it is possible to derive information that has been specified only implicitly (DECKER et al., 1999). Ontology is also a technology that can give support for the development of more intelligent systems (SOWA, 2002). The *intelligent* notion, in our approach, means to use the ontology to assist users to compose more valuable queries, to improve the accuracy of retrieved information, and to extract additional information that is not explicitly stated. All such features are in accord with information-seeking research that states the importance of interactive tools to support users during the search process and that active intervention and interaction are required by the user (MARCHIONINI, 1989; BELKIN, 1993; KUHLTHAU, 1993; INGWERSEN; JÄRVELIN, 2005a; LI; BELKIN, 2010).

The main component of our approach is the ontology. The ontology is the foundation for creating a knowledge base composed of instances. Each instance represents a materialized view of a class defined in the ontology (e.g., given a professor's class, a specific professor is an instance of such a class). In our approach, the user is able to navigate between the instances and concepts of the ontology in a back and forth fashion. This way, the results set has a more active role during the retrieval and browsing process due to the use of the contextual meaning included in each instance. The retrieved information is not presented as simple textual documents but as ontology instances. As ontologies can be composed of more than one class, there can be relationships among classes. These relationships indicate which kind of instances should be retrieved. For instance, if we select the author concept as the central concept, instances of paper and institution concepts are retrieved. The additional information component also uses these relationships, because inference techniques can derive new knowledge building implicit or indirect relationships based on the explicitly defined ones.

Ontologies can become a means of communication between users and the system and help the overcoming of bottlenecks in information access, which is primarily based on keyword searches (SHAH; FININ et al., 2002). In our case, each concept defined in the ontology can be used to compose contextualized queries. Ontology thus can guide users in the formulation of queries and the selection of appropriate terms that are likely to lead to nonempty answer sets. As users are able to explicitly define terms for specific concepts, they are composing more accurate queries automatically. This means that the ontology is also used to disambiguate query terms. Complex queries, using more than one concept, for instance, can also be easily built due to the interface support for choosing concepts and typing the corresponding terms. For instance, using this functionality the user could select the "*author*" concept and type the term "*fonseca*," then select the "*paper*" concept and type the term "*information systems*." Figure 2describes in EBNF the general format of queries applied to a search.

```
search ::=  query  |  contextualized_query
query ::= term [ term ]
contextualized_query ::= concept ":" term [ term ] ";"  [ concept ":" term [ term ] ";" ]
concept ::= <classes described in the ontology>
term ::= letter [ letter ]- |  digit [ digit ]-
letter ::= "a" | "b" |  ...  | "z"
digit ::= "0" | "1" | "2" | "3" | "4" | "5" | "6" | "7" | "8" | "9"
```

Figure 2 - EBNF describing a query in a search

The use of ontologies in information systems gives also an additional possibility that is the capability to use inference techniques to derive additional knowledge. This is possible because the ontology is described in a formal language that includes axioms for specifying relationships between concepts. Hermeneus was conceived to allow the user to take advantage of such inference techniques during the search process. Therefore, besides the ontology's graphical visualization and the list of retrieved instances, users can also visualize and interact with additional information for a given retrieved instance. For example, selecting a retrieved instance about a specific paper, the *additional information component* could show, for instance, similar papers, authors with papers in the same area or papers related to the authors' institutions.

The interface takes advantage of the ontology and the distinct kinds of retrieved information, focusing on a layout that induces easy interaction. Figure 3 shows, on the left side, that the ontology is presented graphically in the *ontology navigator component*. In the middle, the *retrieved instances component* is responsible for presenting the instances that are retrieved according to the users' queries and the chosen *central concept*. Moreover, as instances hold contextual meaning, each retrieved instance can be used to compose contextualized queries automatically. On Figure 3's right side, the *additional information component* is shown. This component depends on the ontology and on the knowledge base. The inference rules use concepts and relationships in the ontology and apply them to the knowledge base. The results are displayed and the user can interact with them and create new or modify existing queries.

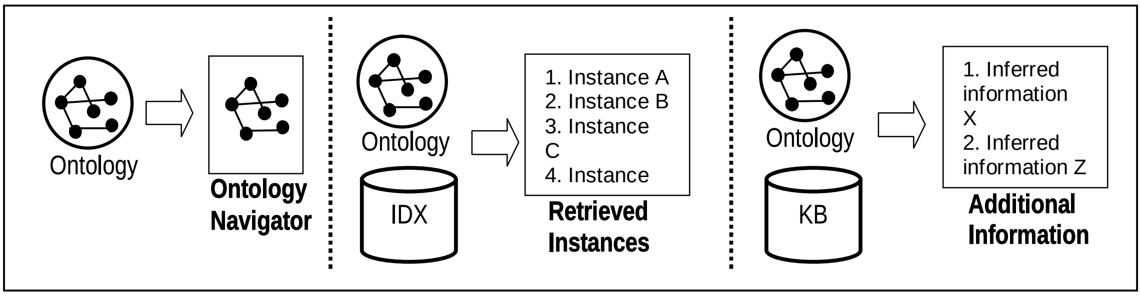

Figure 3 - Ontology-related components available for user's interaction

Furthermore, because the ontology drives the whole process, all the three components (*ontology navigator*, *retrieve instances*, and *additional information*) are strongly integrated. Therefore, actions in a component may affect the others. For example, if the user selects a piece of content from a retrieved instance in the *retrieved instances component* to refine a query, the *ontology navigator component*highlights automatically the concept associated to such content. In addition, the user can select a new retrieved instance in the *retrieved instance component* so as to visualize additional information; thus, the *additional information component* interprets the selected instance and extracts new information based on the ontology and the instances directly related.

One of the main focuses of our system is to help users understand and develop their own information needs. In the following we discuss how our system can help users during the process and how the tools available provide support for users to develop and extend their state of knowledge about a specific information need.

4 DISCUSSION

Although it may seem awkward that users initially do not know exactly what they are looking for, the literature deals with this uncertainty called knowledge gap or state of incompleteness. It is this gap that drives users to use intermediaries, such as IR systems, to supply their information needs. Users, however, have to develop their own information needs first so as to be able to figure out what they are looking for. In this process of trying to understand their own needs, the users' cognitive state may contain doubts or problems in interpretation and then fall into a state of uncertainty (INGWERSEN; JÄRVELIN, 2005a). Users normally have a lack of knowledge, or an anomalous state of knowledge, about the subject they are looking for (BELKIN, 1993). Both the process of understanding the lack of knowledge and the searching for information are cognitive tasks narrowly related to each other during the whole cycle of information acquisition. It reinforces the fact that IR systems should allow users to interact with tools that can affect their state of uncertainty in order to stimulate a transformation in their state of knowledge (MARCHIONINI, 1989; KUHLTHAU, 1993; HERT, 1997; INGWERSEN; JÄRVELIN, 2005a; AZZOPARDI, 2011). Hermeneus was designed with the purpose of giving the users the necessary support for them to develop and understand their own information needs. The hermeneutic circle is the principle behind the system.

4.1 Hermeneus as the translator of the user's information needs into meaningful queries

When interacting with an IR system the most common way to express information needs is composing queries. Nevertheless, formulating precise queries is not an easy task. We tried to address these issues allowing users to compose contextualized queries. Since users are able to interact with the ontology, they can identify which concepts are available and use them to formulate queries. A contextualized query basically is composed of the concept's name and the desired search terms. In our prototype, users can interact with the ontology through mouse-clicks; when they click on a concept it is automatically added to the current query's content so as to assist the composition of contextualized queries. For example, a search for the term "*datamodeling*" (i.e., initially the query contains just the "*datamodeling*" term) combined with a click on the concept *author* available in the ontology, adds the description of such a concept to the query with the colon symbol (i.e., the query now consists of "*data modeling author:*" content) – so the user may just insert the term specifically related to the *author* concept without needing to type the concept's name (e.g., "*data modeling author:Fonseca*"). In developing our solution we followed the hermeneutic circle principle that through the interaction with the ontology, users get insights about their needs and express themselves with more useful queries, and thus obtain more precise results.

The user's state of knowledge is dynamic rather than static, changing as the user proceeds during the searching process (KUHLTHAU, 1991; HERT, 1997; LI; BELKIN, 2010). In our approach, since we retrieve instances and present them in a highly contextualized environment, and each piece of retrieved information has a semantic link with concepts in the ontology, users can improve their queries by adding contextualized content from relevant instances automatically. For example, in a very common situation where a user types just a few terms (e.g., "*semantic*") to express his or her need, the retrieved instances can help him or her to find a more precise description and use such description dynamically. For instance, in our prototype when searching for the term "*semantic*," users can visualize and interact with the following terms related to the *keyword* concept: *latent semantic indexing, semantic network*, *semantic information*, *semantic-based information retrieval*, *semantic-based web retrieval systems*, *semantic web*, and *semantic search*. With this kind of exploratory interaction users are shielded from details of query formulation besides being able to construct more useful search statements (CAPRA III; PÉREZ-QUIÑONES, 2005; LI; BELKIN, 2010). For example, searching for the term "*web*" in our prototype, the user receives 169 retrieved instances. Next, clicking on the "*semantic web*" content present in one of the retrieved instances leads to a new query composed of "*web*" term and "*keyword: semantic web;*" content. The result is only 11 instances. Therefore, using our system users can experiment with retrieved content to formulate new queries as part of the process of understanding and developing their information needs.

The ontology can also be used to define which kind of retrieved instances should be shown through the selection of a *central concept*. This functionality enables users to have different visualizations of retrieved instances for the same query. Ontologies can be used to derive further views of knowledge, particularly in the exploration of new navigation purposes (STAAB et al., 2001). Since the process of finding information is dynamic and constantly changing (INGWERSEN; JÄRVELIN, 2005a), visualizing retrieved information with different perspectives may help users get insights about their needs. In our prototype, for example, when searching for "*information science*," there are 58 retrieved instances grouped by the *paper* concept (i.e., the *central concept* is *paper*). After changing the visualization of the results set to be grouped by the *author* concept (i.e., the *central concept* was changed to *author*) there are then 73 instances of authors. In this case, the user was able to visualize the results set with authors and their respective information (i.e., instances of papers and institutions for each retrieved author). Such functionality is in accordance with Vakkari's (2003 p. 426) theory, who says that "modeling and developing systems features, like features of interfaces that support users in articulating their information needs by using terms from potentially relevant documents grouped according to the concepts in the users' queries reflecting the task, would enhance retrieval performance."

4.2 Extracting and interacting with additional information

Additional information helps users have a better understanding about the domain and facilitates them to learn more about a specific instance (GOLBREICH, 2004; MURDOCK et al., 2005). In our approach, the *additional information component* enables users to visualize information that is not explicitly stated in the result list of retrieved instances. Therefore, inferring new information for each retrieved instance enables users to get insights about their information needs.

Inference, according to Decker et al. (1999), is the logical process of derivation of new data from a collection of data. We use this technique combined with rules to extract information that was only implicit before. In our prototype, for example, when the retrieved instances are grouped by the *paper* concept (i.e., the *central concept* is *paper*) and users pick out a specific instance, the *additional information component* presents inferred information about authors and their institutions. In this case, as shown in Table 2, it is presented the authors' most frequent keywords, considering all papers stored in the knowledge base, the institution to which the authors are affiliated with, and the authors' collaborators. In addition, users can see the most frequent keywords for each institution and the papers that mention such institution (Table 2). The information available in the *additional information component* can also be used for additional exploratory activity, because users can interact with such information and refine their searches dynamically. This interactive process has the same principle available in the *retrieved instances component*.

Table 2 - A sample of the additional information presented in the *additional information component*

Retrieved instance (paper)	Additional Information
Title: The double role of ontologies in information science research Author: Frederico Fonseca Keywords: information systems; ontology; ontologies; engineering; science; information science; ontology-driven information science;	Frederico Fonseca is connected to Pennsylvania State University, University Park, PA 16802, his most frequent keywords are: ontologies, ontology, information science, engineering, science; and she or he cooperates with: - Gilberto Câmara; - James E. Martin; Pennsylvania State University, University Park, PA 16802 has 18 paper(s), written by 36 author(s), and its most frequent keywords are: web, web search, web search engine, information retrieval, ontologies.

4.3 Ontology graphical interaction

The aim of a graphical tool is to help human beings gain insight into data. Visualizing an ontology means to view its concepts and respective relationships (SPENCE, 2000). The added value of an ontology's graphical presentation lies in its expressivity because concepts and their relationships (the vocabulary of the domain) become easier to detect (FLUIT et al., 2002; MELGAR et al., 2010). Our approach takes advantage of this and enables users' interaction with the ontology during the whole process of searching for information. The interface of the *ontology navigator component* has been designed to stimulate the user to explore the information shown. It can assist users in understanding their knowledge gap and then specify their information needs more precisely.

Using Hermeneus, users can identify visually which concepts were used in a search. This functionality highlights all concepts belonging to a contextualized query with a distinct color.

For example, searching for "*author: Smith*," the concept *author* has a different color (yellow) in the ontology. Users can also view what information is specifically related to each concept in accordance with the query. Considering the prior example, the *author* box in the ontology shows "*Smith*" as its content. These kinds of graphical features help users to contextualize their needs at a time that their understanding of the problem is still developing.

The graphical tool also enables users to have different perspectives of the results set. This functionality has already been described before, but it is important to emphasize that the changes in the results set's perspective is done through mouse-clicks on the concepts available graphically. The type of information the results set is showing is related to the *central concept*. The concept used as the *central concept* is drawn with a different shape (square) in order to facilitate visual identification of which type of information is being presented.

In what follows, we describe some approaches related to information seeking research and ontology applied in the context of information retrieval systems which were relevant to our proposal.

5 RELATED WORK

Although there are several works describing IR approaches in the context of information seeking, none of them addresses all the points we do. Spink et al. (2002) described a theoretical framework to analyze the actions taken by information-seekers in their search for information. Situated actions, for instance, according to the authors, occur during IR interactions over time due to judgments made by seekers during an evolving information-seeking process. Other concepts were also described like uncertainty, cognitive styles, and successive searching behavior. Some variables discussed by Spink et al. (2002) were also discussed in our approach, but with a different perspective. In our approach we opted to use the hermeneutic circle to deal with situated actions. The hermeneutic circle explains that seekers have the fore-structure, or prejudice, which is the initial idea the user has about the subject, and how user might come to new insights when they interact with new information in a continual back and forth movement. The hermeneutic circle along with the ontology were used to create our framework for it to be an intermediary between the user and his/her information needs.

Cole (1998) presented a theoretical framework for intelligent IR on diagnostic procedures used in medicine, called differential diagnosis. According to the author, an IR system can be an intermediary that interacts with users like doctors do with their patients. In another work Cole et al. (1998) described a prototype based on the theoretical framework. In such a prototype, the user inserted an assignment that broadly described his/her information needs. The system applied an uncertainty measure on the user's report and then assigned a class in a classification scheme of information needs. Afterward, the system visualized and summarized the information need to the user, so that he/she could "react to it in an intuitive, easy to grasp manner" (COLE et al., 1998, p. 722). While Cole's system requires users to describe their information needs without helping, our approach enables the user to take advantage of domain knowledge so as to get insights about his/her information needs and to compose queries more easily. Furthermore, in our approach the user can interact with retrieved information in order, for example, to redefine his/her search dynamically.

There are also several works that apply ontology to enhance IR approaches. We use ontology to materialize the idea of hermeneutic circle in the context of information seeking and retrieval. McGuinness (1999), for instance, developed the initial idea of ontology driven IR. Her work presented an ontology-enhanced online search for medical documents. Users could combine a simple keyword search with the choice of one of the content areas displayed in the same window. The content areas were presented in a tree structure as a taxonomy where users were able to navigate and choose one of them. This approach used a hierarchical structure and limited the composition of more complex queries relating concepts from different parts of the hierarchy. To link one keyword to a specific concept was also not possible (i.e., the typed keywords were applied to all selected categories). Furthermore, the results screen showed only the retrieved records and did not present the concepts associated with the information used in the search. Thus it is harder for the user to contextualize the results in the domain knowledge. In our approach, users are allowed to use the ontology as a guide to compose queries and define explicitly for each concept its specific terms (e.g., *author: fonseca paper: ontology*). In addition, in the results screen users can still interact with the ontology and see which concepts were used in the current search, besides being able to interact with the retrieved instances to narrow down the scope of their searches more dynamically.

6 CONCLUSIONS AND FUTURE WORK

If the user does not know exactly what she or he wants, how can a framework for information seeking and retrieval help in this process? We looked into the philosophy of Heidegger and Gadamer to find the principles to guide our solution. We used the concept of the hermeneutic circle as the support for the information seeking process. In the hermeneutic circle users develop their ideas as while browsing the information and the concepts that represent the information.

The solution we presented here was the framework called Hermeneus. We choose the name Hermeneus, which in Greek means the interpreter or translator (KINGSLEY, 1993), following Kuhlthau's (1991, 1993) concept that an IR system is an intermediary between the user's need for information and the achievement of the resolution of that need. Information seeking theory helped us understand that the user requires an environment where he or she can intervene and interact more actively. Finding information is an exploratory activity where users apply their knowledge and intuition combined with strategies and tools to find the desired information (CAPRA III; PÉREZ-QUIÑONES, 2005; LI; BELKIN, 2010). Our framework enabled users to engage in a sequence of interactions that helped them understand what they were looking for.

Hermeneus is based on ontologies in order to enhance the interactivity with the user. We used a dynamic process for query building so that users could browse the ontology concepts and the instances of these concepts in a back and forth way. Doing so users could also change modes of visualization based on each concept described in the ontology. Users could also interact with retrieved instances so as to use their contents to begin new searches automatically. Furthermore, additional information could be extracted for each retrieved instance. Users were able to visualize and interact with knowledge which before was just implicit stated in the ontology and in the knowledge base.

Hermeneus addressed Järvelin's (2004) and Ingwersen's (2005b) requisites of a search activity that is easier, faster, and intuitive. In terms of easiness, our approach allowed users to access the domain knowledge through keywords search and to interact with the ontology graphically – therefore they could compose queries using terms according to their initial state of knowledge and get insights when navigating in the ontology concepts. Regarding fastness, because users' state of knowledge can be influenced by the retrieved information, our approach enabled users to use query refinement as an assistant to compose contextualized queries dynamically and to extract additional information for each retrieved instance. Such operations are in accordance with Ingwersen and Järvelin (2005b) theory about information seeking who states that relevance is a cognitive, situational, and dynamic phenomenon. Finally, concerning intuitiveness, users could use concepts described in the ontology to compose contextualized queries, use the retrieved content to redefine their searches, and use the ontology to visualize the results set in different perspectives. These functionalities, according to Card et al. (1999) and Albertoni et al. (2005), augment user's cognition and thus reach the goal state of resolution more quickly.

As future work, we are looking into a new way to present the retrieved information taking advantage of the semantic content described in the instances and the ontology. We think that a more graphical way to show the results set can enrich even more the interactivity between the users and the system. We also understand that users' traceability could be stored so as to help other users when they are searching for similar information. In this case, the system could "understand" new users' behavior *on-the-fly* enhancing the assistance throughout the seeking process and giving new clues about their searches. We also envisage the creation of a mechanism that could be used by users to define their own ontologies dynamically and thus configure the IR system according to their conceptualizations of a specific domain. User, in this case, will be able to create a personalized IR system and share it with others if she or he wants.

A ABORDAGEM FILOSÓFICA PARA BUSCA E RECUPERAÇÃO DE INFORMAÇÃO INTITULADA HERMENEUS

Resumo

Pessoas que buscam por informações sempre têm uma pergunta em mente, mesmo quando não sabem exatamente o que é. De que forma um sistema de recuperação de informação pode dar suporte a usuários enquanto eles ainda estão moldando as suas necessidades por informação? Recorreu-se ao conceito filosófico do círculo hermenêutico de Heidegger e Gadamer para fornecer os princípios de um sistema desse tipo. Na presente proposta os usuários são capazes de desenvolver as suas idéias enquanto navegam nas informações e nos conceitos que representam tais informações. Utilizou-se ontologias para desenvolver esta abordagem hermenêutica. Ontologias podem melhorar os sistemas de recuperação de informação quanto à sua recuperação e apresentação da informação, o que torna a tarefa de encontrar informação mais eficaz, eficiente e interativa. Hermeneus é o nome dado ao framework proposto, que em grego significa o intérprete ou tradutor. Hermeneus funciona como um intermediário que auxilia o usuário a passar do estado inicial de necessidade por informação para o estado de solução. O sistema proposto pretende ser a ponte entre a questão do usuário e a resposta a ser encontrada enquanto ele navega nos conceitos da ontologia e nas instâncias desses conceitos em forma de vai e vem.

Palavras-chave: *Recuperação e busca de informação. Ontologia. Círculo hermenêutico. Hermeneus.*

Artigo recebido em 26/03/2011 e aceito para publicação em 24/08/2011

REFERENCES

ALBERTONI, R. et al. Information search: the challenge of integrating information visualization and semantic web. **Sixteenth Int. Workshop on Database and Expert Systems Applications**, p. 529-533. 2005.

AZZOPARDI, L. The economics in interactive information retrieval. In: ANNUAL INTERNATIONAL ACM SIGIR CONFERENCE, 34, Beijing, China: ACM 2011.

BELKIN, N. J. **Interaction with Texts**: Information Retrieval as Information-Seeking Behavior. Information Retrieval '93: Von derModellierung zu Anwendung, 1993.

______; CROFT, W. B. Information filtering and information retrieval: two sides of the same coin? **Communications of the ACM**, v. 35, n. 12, p.29-38. 1992.

BEPPLER, F. D. **Um Modelo para Recuperação e Busca de Informação Baseado em Ontologia e no Círculo Hermenêutico**. Florianópolis: UFSC, 2008. 123 f. Tese (Doutorado em Engenharia e Conhecimento). Universidade Federal de Santa Catarina. 2008.

______; FONSECA, F. T.; PACHECO, R. C. S. Hermeneus: uma arquitetura para recuperação e Busca de Informação. **DataGramaZero** - Revista de Ciência da Informação, v. 10, n. 2, p. 1-25, 2009.

BERNSTEIN, R. J. Beyond **Objectivism and relativism:** science, hermeneutics, and praxis. Philadelphia, PA, USA: University of Pennsylvania Press. 1983

CAPRA III, R. G.; PÉREZ-QUIÑONES, M. A. Using web search engines to find and refind information. **IEEE Computer Society**, v. 38, n. 10, p. 36-42. 2005.

CARD, S. K. et al. **Readings in information visualization:** using vision to think. [S.l]: Morgan Kaufmann, 1999

COLE, C. Intelligent information retrieval: diagnosing information need. Part I. The Theoretical Framework for Developing an Intelligent IR Tool. **Information Processing and Management**, v. 34, n. 6, p. 709-720, 1998.

COLE, C. et al. Intelligent information retrieval: diagnosing information need. Part II. Uncertainty Expansion in a Prototype of a Diagnostic IR Tool. **Information Processing and Management**, v. 34, n. 6, p. 721-737, 1998.

DECKER, S. et al. **Ontobroker:** ontology based access to distributed and semi-structured information. DS-8: Kluwer Academic Publisher, 1999.

FONSECA, F. The double role of ontologies in information science. **Journal of the American Society for Information Science and Technology**, v. 58, n. 6, p. 786-793. 2007

FLUIT, C. et al. Ontology-Based Information Visualization. In: (Ed.). **Visualising the Semantic Web:** Springer Verlag, [S.l], [S. n], 2002.

GADAMER, H. G. **Truth and Method.** (Original publication 1960, as Wahrheit und Methode). Crossroad, 1989

GEERTZ, C. From the Native's Point of View: On the Nature of Anthropological Understanding. In: RABINOW, P.; SULLIVAN, W. M. (Ed.). **Interpretive social science**: a reader. Berkeley: University of California Press, 1979. p. 225-241

GOLBREICH, C. **Combining rule and ontology reasoners for the semantic web**. RuleML. Hiroshima, Japan: Springer-Verlag, 2004.

GUARINO, N. **Formal Ontology in Information Systems**. Netherlands: IOS Press, 1998.

HEIDEGGER, M. **On the way to language**. San Francisco: Harper & Row. 1982.

______; Stambaugh, J. **Being and time:** A translation of sein und zeit. Albany, NY: State University of New York Press. 1996.

HERT, C. A. **Understanding information retrieval interactions:** theoretical and practical implications. London: Ablex Publishing Corporation, 1997

INGWERSEN, P.; JÄRVELIN, K. The sense of information: understanding the cognitive conditional information concept in relation to information acquisition. **Lecture Notes in Computer Science, Springer**, v. 3507, p. 7-19, 2005a.

______. **The Turn:** Integration of Information Seeking and Retrieval in Context. Netherlands: Springer, 2005b (The Information Retrieval Series).

Ingwersen, P. et al. Panel: Perspectives on Adaptivity in Information Retrieval Interaction (PAIRI). **ASIST**. 2010.

JANSEN, B. J. et al. Determining the Informational, Navigational, and Transactional intent of Web Queries. **Information Processing and Management**, v. 44, p. 1251-1266, 2008.

JÄRVELIN, K.; Ingwersen, P. Information Seeking Research Needs Extension Towards Tasks and Technology. **Information Research:** An International Electronic Journal, v. 10, n. 1, 2004.

KINGSLEY, P. Poimandres: the etymology of the name and the origins of the hermetica. **Journal of the Warburg and Courtauld Institutes**, v. 56, p. 1-24, 1993.

KUHLTHAU, C. C. Inside the search process: information seeking from the user's perspective. **Journal of the American Society for Information Science**, v. 42, n. 5, p. 361-371, 1991.

______. **Seeking Meaning:** A Process Approach to Library and Information Services. USA: Ablex Publishing Corporation, 1993

LI, Y.; BELKIN, N. J. An exploration of the relationship between work task and interactive information search behavior. **Journal of the American Society for Information Science and Technology**, v. 61, n. 9, p. 1771-1789, 2010.

MARCHIONINI, G. Information-seeking strategies of novices using a full-text electronic encyclopedia. **Journal of the American Society for Information Science**, v. 40, n. 1, p. 54-66, 1989.

McGUINNESS, D. L. **Ontology-enhanced search for primary care medical literature**. Phoenix, Arizona: Int. Medical Informatics Association, 1999.

MELGAR, H. A. et al. Knowledge Retrieval in the Anatomical Domain. IHI. Arlington, Virginia, USA: ACM, 2010.

MURDOCK, J. W. et al. **Encoding Extraction as Inferences**. AAAI Spring Symposium on Metacognition on Computation: AAAI Press, Stanford University, 2005.

SHAH, U. et al. Information Retrieval On The Semantic Web. In: INTERNATIONAL CONFERENCE ON INFORMATION AND KNOWLEDGE MANAGEMENT, 11, 2002. Virginia, USA, 2002. p. 461-468.

SILVERSTEIN, C. et al. Analysis of a Very Large Web Search Engine Query Log. **ACM SIGIR Forum**, v. 33, n. 1, p. 6-12, 1999.

SONNENWALD, D. H. et al. A Research Method to Investigate Information Seeking Using the Concept of Information Horizons: An Example from a Study of Lower Socio-Economic Students' Information Seeking Behaviour. **The New Review of Information Behaviour Research**, v. 2, p. 65-86, nov. 2001.

SOWA, J. F. Architectures for intelligent systems. **Systems Journal, Artificial Intelligence**, v. 41, n. 2, p. 331, 2002.

SPENCE, R. **Information visualization**. [S.l]: Addison Wesley, 2000.

SPINK, A. et al. Information-seeking and mediated searching. Part 1. Theoretical Framework and Research Design. **Journal of the American Society for Information Science**, v. 53, n. 9, p. 695-703, 2002.

Spink, A. et al. Searching the web: the public and their queries. **Journal of the American Society for Information Science**, v. 53, n. 2, p. 226-234, 2001.

STAAB, S. et al. Knowledge processes and ontologies. **Intelligent Systems, IEEE**, v. 16, n. 1, p. 26-34, 2001.

VAKKARI, P. Task-based information search. **Annual Review of Information Science and Technology**, v. 37, p. 413-464, 2003.