KNOWLEDGE DISCOVERY AND EXCHANGE

Towards a Web-based Application for Discovery and Exchange of Revealed Knowledge

S. J. Overbeek
e-office B.V., Duwboot 20, 3991 CD Houten, The Netherlands, EU
Sietse.Overbeek@e-office.com

P. van Bommel, H. A. (Erik) Proper, D. B. B. Rijsenbrij
Institute for Computing and Information Sciences, Radboud University Nijmegen
Toernooiveld 1, 6525 ED Nijmegen, The Netherlands, EU
P.vanBommel@cs.ru.nl, E.Proper@cs.ru.nl, D.Rijsenbrij@cs.ru.nl

Keywords: Knowledge discovery, knowledge exchange, knowledge market.

Abstract: Web technologies enable the discovery and exchange of knowledge from many different locations using many different channels. This implies that one is able to discover and exchange knowledge while using a PDA when traveling by train for instance. A provisional Web-based application referred to as 'DEXAR: *Discovery and eXchange of Revealed knowledge*' is therefore introduced to illustrate the possibilities of the Web in the process of knowledge discovery and exchange. This is illustrated by an example from the medical domain. Before focussing on this Web application however a better understanding of knowledge discovery and exchange is needed to be able to determine what kind of Web-based support is desired and feasible. Thus, a knowledge market paradigm and a knowledge discovery paradigm are discussed in detail.

1 INTRODUCTION

Knowledge discovery and exchange might be associated directly to (Web) technologies such as search engines, agent technology, mining tools, meta-data standards, query languages, query protocols, etcetera. Obviously, Web technologies can assist in discovering and exchanging knowledge from many different *locations* such as at home, at the office or at the university. Web technologies can also assist in discovering and exchanging knowledge using a variety of *channels*, such as a desktop computer, a notebook and a PDA. Before focussing on a possible Web-based application however a better understanding of knowledge discovery and exchange is needed to be able to determine what kind of Web-based support is desired and feasible.

The latter can be acquired by using several reference models (in terms of a knowledge market paradigm and a knowledge discovery paradigm respectively) which depict essential knowledge market and knowledge discovery mechanisms on a conceptual level. The actual knowledge discovery mechanism described only focuses on the discovery of revealed knowledge. This is knowledge which is indeed known to an individual or the organization. The actual discovery of concealed knowledge is not elaborated in depth, because it will probably fit more in a knowledge *mining* paradigm than in the knowledge discovery paradigm introduced in this paper.

Based on the reference models a provisional Web-based application is elaborated, together with an initial user interface to assist the user in discovering and exchanging revealed knowledge. The application is used in a medical context, in that an assistant radiologist acquires medical knowledge utilizing knowledge discovery and exchange mechanisms.

In section 2 the fundamentals of knowledge discovery and exchange are elaborated, because it is necessary to understand *what* is going to be discovered and exchanged. A knowledge market paradigm is then described in section 3, followed by a knowledge discovery paradigm in section 4 which is a specialization of the knowledge market paradigm. After the theory has been discussed, a provisional Web-based application implements the knowledge market and discovery paradigms into practice in section 5. Section 6 briefly compares our reference models with other approaches in the field and outlines the benefits of our approach compared to others. Section 7 concludes this paper.

2 FUNDAMENTALS OF KNOWLEDGE DISCOVERY AND EXCHANGE

Exploring the fundamentals of knowledge is necessary to gain a better understanding of that what we would like to discover and exchange. Before elaborating on the notion of knowledge, it is relevant to mention that it can be regarded as ‘wrapped’ in information, whilst information is ‘carried’ by data (expressions in a symbol language) (Liang, 1994).

To determine possible knowledge types which can be discovered and exchanged, implicit knowledge and explicit knowledge can be elaborated at first. (Nonaka and Takeuchi, 1995) distinguish implicit knowledge and explicit knowledge. *Implicit knowledge* comprises knowledge which is implicitly present in people’s heads, such as skills which are difficult to make explicit. The way a physician makes a decision for specific treatment is related to such skills. Implicit knowledge is closely related to what is generally experienced as intuition. *Explicit knowledge* comprises knowledge which can be expressed in terms of facts, rules, specifications or textual descriptions. The difference between implicit and explicit knowledge is relevant with regard to Web-based knowledge discovery and exchange, because the discovery and exchange of implicit knowledge will require different forms of Web-based support than the discovery and exchange of explicit knowledge.

Besides discerning implicit and explicit knowledge another relevant distinction can be made. Sometimes knowledge is present while one is not aware of that knowledge. This varies from hidden skills of workers (for an individual or for the organization), via knowledge which is present in documents but not properly indexed, to knowledge which is hidden in undiscovered patterns in data collections (the basis for data mining). (Hoppenbrouwers and Proper, 1999) already introduced this difference between revealed and concealed knowledge. In contrast to implicit and explicit knowledge the knowledge *status* is considered instead of the fundamental knowledge *type*.

The difference between *implicit* and *explicit* on the one hand and *revealed* and *concealed* on the other hand can be depicted in a 2x2 matrix as is shown in figure 1. The following combinations are then possible:

- Implicit & concealed knowledge: e.g. competencies or expertise of a worker unknown to the organization;
- Explicit & concealed knowledge: e.g. valuable insights concealed in available data collections (to be discovered by data mining);
- Implicit & revealed knowledge: e.g. known expertise of a worker which can be appealed to;
- Explicit & revealed knowledge: e.g. best-practice documentation, knowledge bases, scientific papers, etcetera.

	Implicit	Explicit
Concealed	Unknown competences	Unknown patterns and structures in data
Revealed	Known competences	Documented knowledge

Figure 1: Four knowledge types, adapted from (Hoppenbrouwers and Proper, 1999).

The four knowledge types require different forms of Web-based support, varying from support which is aimed at expertise analysis, pro-active knowledge exchange, human resource related information systems, planning systems, data mining and documentation management. In the broad sense, Web-based support can be applied to all knowledge types, nevertheless Web-based support is usually applied more easily to explicit knowledge than implicit knowledge.

At this point we can aim on the actual *knowledge discovery and exchange*, which is illustrated by the reference models together with a provisional Web-based application from the medical domain. Before focussing on discovery and exchange of revealed knowledge only (in a knowledge discovery paradigm), a more general knowledge market paradigm is discussed.

3 THE KNOWLEDGE MARKET PARADIGM

Several knowledge types have been discussed up till now. In practice the difference between concealed and revealed knowledge is especially of importance. Revealed knowledge can be localized (even when it is implicit), but concealed knowledge can not be localized (even when it has been made explicit in the past).

These knowledge types materialize in a *knowledge market paradigm* as is depicted in figure 2.

3.1 Roles in the Knowledge Market

As in every market model the sketched knowledge market contains the following roles: a *broker* role, a

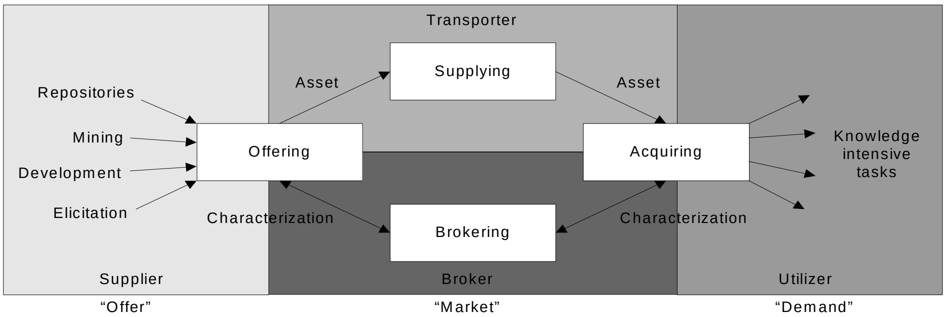

Figure 2: A knowledge market paradigm.

supplier role, a *transporter* role and a *utilizer* role. Formally, this set of roles can be represented as:

$$\mathcal{RO} \triangleq \{\texttt{broker}, \texttt{supplier}, \texttt{transporter}, \texttt{utilizer}\} \quad (1)$$

The four roles of the knowledge market paradigm wish to achieve specific objectives within the market and enable to abstract from the actors that will eventually enact the role. A role enactment is a specific *fulfilment* of such a role by any eligible entity, expressed by the function $\mathsf{Enact} : \mathcal{RE} \rightarrow \mathcal{RO}$, where $\mathcal{RE}$ is the set of all role enactments within the domain of the knowledge market. Given the role enactment e of a role $\mathsf{Enact}(e)$, we can view the actor that specifically enacts the role as a function $\mathsf{Player} : \mathcal{RE} \rightarrow \mathcal{AC}$. Here $\mathcal{AC}$ represents the specific set of actors. Each role is discussed briefly in this section, also illustrated by a practical example. The following text about role enactments incorporates ideas from (Gils et al., 2006) to illustrate the essence of role enactments in the knowledge market. Since an enactment indicates an actor 'in a role' we know that an actor and a role combination uniquely determines an enactment:

$$\mathsf{Player}(e_1) = \mathsf{Player}(e_2) \wedge \mathsf{Enact}(e_1) = \mathsf{Enact}(e_2) \Rightarrow e_1 = e_2 \quad (2)$$

For an enactment $e \in \mathcal{RE}$ the following notation is introduced:

$$\langle a, r \rangle \triangleq e \text{ such that } \mathsf{Player}(e) = a \wedge \mathsf{Enact}(e) = r \quad (3)$$

This can be illustrated by the following example. Let an *assistant radiologist* denoted by a be an actor that can play two roles. He either plays the role of type *broker* denoted by r_1, or the role of type *utilizer* denoted by r_2. Both $e_1 = \langle a, r_1 \rangle$ and $e_2 = \langle a, r_2 \rangle$ are enactments such that $\mathsf{Player}(e_1) = a$, $\mathsf{Player}(e_2) = a$, $\mathsf{Enact}(e_1) = r_1$ and $\mathsf{Enact}(e_1) = r_2$.

3.2 Tasks in the Knowledge Market

An actor in the knowledge market executes several tasks while playing a role. The utilizer for instance 'utilizes' knowledge by executing *knowledge intensive tasks* as is depicted in figure 2. Actors in the knowledge market execute different tasks depending on the role they play. A specific fulfilment of such a task is expressed by $\mathsf{Task} : \mathcal{TI} \rightarrow \mathcal{TA}$, where $\mathcal{TI}$ is the set of tasks which are executed by an actor (referred to as task instances) and $\mathcal{TA}$ is the set of tasks. Given the task instantiation i of a task $\mathsf{Task}(i)$, we can view the actor that is enacting a role and also specifically executing a task as a function $\mathsf{Exec} : \mathcal{TI} \rightarrow \mathcal{RE}$. The essence of task execution in the knowledge market can be illustrated in the same way as in the role enactment part of section 3.1. Since an execution denotes an enactment in a task we know that an enactment and a task combination uniquely determines an execution:

$$\mathsf{Exec}(i_1) = \mathsf{Exec}(i_2) \wedge \mathsf{Task}(i_1) = \mathsf{Task}(i_2) \Rightarrow i_1 = i_2 \quad (4)$$

For a task instance $i \in \mathcal{TI}$ the following notation is introduced:

$$\langle e, t \rangle \triangleq i \text{ such that } \mathsf{Exec}(i) = e \wedge \mathsf{Task}(i) = t \quad (5)$$

This can also be illustrated by an example. Assume that a *radiologist* enacts a 'utilizer' role. This enactment is denoted by e. While enacting the role he can execute a task of type *patient monitoring* denoted by t_1 and a task *patient diagnosing* denoted by t_2. Both $i_1 = \langle e, t_1 \rangle$ and $i_2 = \langle e, t_2 \rangle$ are executions such that $\mathsf{Exec}(i_1) = e$, $\mathsf{Exec}(i_2) = e$, $\mathsf{Task}(i_1) = t_1$ and $\mathsf{Task}(i_1) = t_2$.

Now that roles and tasks in the knowledge market have been discussed on a conceptual level, the function of the knowledge market in terms of supply and demand is elaborated in the next section.

3.3 Supply and Demand

The 'merchandize' within the paradigm consists of knowledge assets $\mathcal{KA}$. These *assets* are tradeable forms of revealed knowledge, which are transported physically by the transporter. In the context of a Web-based application, a transporter role can be enacted by an actor equalling a Web protocol on the application layer according to the OSI model (ISO/IEC, 1994), such as FTP, GOPHER, HTTP, IRC or TELNET.

Knowledge assets are not necessarily *explicit* knowledge assets, though. Implicit knowledge inside people's heads is also tradeable, because one can take its implicit knowledge to a certain situation where that implicit knowledge is wanted. This is what e.g. physicians do when explaining a patient's status to a colleague. In that case the air functions as a transporter. In the context of the Web however, a knowledge asset can be defined as any entity that is accessible on the Web that can provide knowledge to other entities connected to the Web.

The goal of a supplier is to deliver knowledge, which requires a 'client' who would like to utilize the knowledge. This is only possible if the supplier is able to make clear what is on offer, hence it is vital that the knowledge is correctly *characterized*. This is not always an easy job because terminology issues can throw a spanner in the works. Poor characterizations can inevitably lead to supplying irrelevant knowledge, or omitting to supply relevant knowledge.

On the supply side of the knowledge market various resources can be accessed: repositories, data collections, mining, active knowledge development or questioning experts (elicitation). A reliable supplier offers *revealed knowledge* which is localizable and available. It is possible to offer implicit knowledge, e.g. by means of a reliable expert, as long as is assured that the implicit knowledge can be applied by the utilizer (albeit a certain *competence*).

The potential utilizer is searching for knowledge, but does not know if that knowledge can be found. Often a utilizer does not even know which knowledge is necessary to fulfil the need. The knowledge is concealed for the potential utilizer, but does certainly not have to be concealed for the knowledge supplier. Characterization is key here, which matches the knowledge demand with the knowledge to be supplied. The broker plays a very important role in matching supply and demand. It can be said that knowledge discovery comes into play when potential utilisers do not know beforehand which knowledge is required to fulfil their needs.

Now that we have focussed on supply and demand in the knowledge market, the actual exchange of knowledge (albeit knowledge assets or knowledge which is exchanged in a characterization process) within such a market is elaborated in the next section.

3.4 Levels of Knowledge Exchange

In the knowledge market paradigm, two levels of knowledge exchange can be discerned: *instance* level knowledge and *meta* level knowledge:

$$\mathcal{KL} \triangleq \{\texttt{instance-knowledge}, \texttt{meta-knowledge}\} \tag{6}$$

What is exchanged between the supplier, transporter and utilizer can be classified as *instance* level knowledge. On the instance level, the actual knowledge assets which are *utilized* by an actor who enacts the utilizer role are intended. When knowledge exchange on the instance level is concerned, the knowledge assets are part of the knowledge *output* of the supplier as well as the transporter. The assets are then *input* for the utilizer. Input and output of knowledge (on the instance level) can be represented as $\mathsf{In}, \mathsf{Out} : \mathcal{RE} \to \wp(\mathcal{KA})$, where $\mathsf{In}(e)$ determines the input in terms of knowledge assets of an actor enacting a role e. The function $\mathsf{Out}(e)$ then determines the output. So knowledge exchange on the instance level can be represented as:

$$\begin{aligned} x,y,z \in \mathcal{RE} \land \mathsf{Enact}(x) = \texttt{supplier} \quad &\land \\ \mathsf{Enact}(y) = \texttt{transporter} \quad &\land \\ \mathsf{Enact}(z) = \texttt{utilizer} \quad &\Rightarrow \\ \mathsf{In}(y) \cap \mathsf{Out}(x) \neq \emptyset \lor \mathsf{In}(z) \cap \mathsf{Out}(y) \neq \emptyset & \end{aligned} \tag{7}$$

The intersection of knowledge input and output is however considered as an *empty set* when only a single actor is enacting a role:

$$\forall_{e \in \mathcal{RE}} [\mathsf{In}(e) \cap \mathsf{Out}(e) = \emptyset \land \mathsf{MIn}(e) \cap \mathsf{MOut}(e) = \emptyset] \tag{8}$$

A *meta* level of knowledge exchange always contains a formulation in terms of a question or a query which reasons about knowledge which a utilizer wants to receive. Meta level knowledge comprises the knowledge which is exchanged between the utilizer, the broker and the supplier in the process of *matching supply and demand*. Just that knowledge which flows between the aforementioned three roles in the characterization process is intended here. This knowledge is dubbed *characterization knowledge*, which is represented by the set $\mathcal{CK}$. Input and output of knowledge (on the meta level) can be represented as $\mathsf{MIn}, \mathsf{MOut} : \mathcal{RE} \to \wp(\mathcal{CK})$, where $\mathsf{MIn}(e)$ determines the input in terms of characterization knowledge of an actor enacting a role e. The function $\mathsf{MOut}(e)$ then determines the output. So knowledge exchange on the meta level can be represented as:

$$\begin{aligned} x,y,z \in \mathcal{RE} \land \mathsf{Enact}(x) = \texttt{utilizer} \quad &\land \\ \mathsf{Enact}(y) = \texttt{broker} \quad &\land \\ \mathsf{Enact}(z) = \texttt{supplier} \quad &\Rightarrow \\ \mathsf{MIn}(y) \cap \mathsf{MOut}(x) \neq \emptyset \quad &\land \\ \mathsf{MIn}(x) \cap \mathsf{MOut}(y) \neq \emptyset \quad &\lor \\ \mathsf{MIn}(z) \cap \mathsf{MOut}(y) \neq \emptyset \quad &\land \\ \mathsf{MIn}(y) \cap \mathsf{MOut}(z) \neq \emptyset & \end{aligned} \tag{9}$$

In the following section, a scenario is described in which the knowledge market paradigm materializes as a whole.

3.5 A Knowledge Market Scenario in the Medical Domain

The knowledge market paradigm comes to life when it is illustrated by a practical problem from the medical domain. Assume that an assistant radiologist denoted by actor a_1 requires knowledge about pneumonia when executing a knowledge intensive task. This task is referred to as *patient diagnosing* denoted by t_1 while enacting the role of *utilizer* denoted by u. Formally, $e_1 = \langle a_1, u \rangle$ is an enactment such that $\mathsf{Player}(e_1) = a_1$ and $\mathsf{Enact}(e_1) = u$. Also, $i_1 = \langle e_1, t_1 \rangle$ is a task execution such that $\mathsf{Exec}(i_1) = e_1$ and $\mathsf{Task}(i_1) = t_1$.

The assistant radiologist finds a senior radiologist denoted by a_2. This actor plays the role of *broker* denoted by b. Let $e_2 = \langle a_2, b \rangle$ be an enactment such that $\mathsf{Player}(e_2) = a_2$ and $\mathsf{Enact}(e_2) = b$. Actor a_1 explains his specific knowledge need expressed by $\mathsf{MIn}(e_2) \cap \mathsf{MOut}(e_1) \neq \emptyset$. To fulfil the task *assistant support* the senior radiologist probably asks some additional questions to understand what the assistant needs which is expressed by $\mathsf{MIn}(e_1) \cap \mathsf{MOut}(e_2) \neq \emptyset$. The broker's task fulfilment can be formalized as follows: $i_2 = \langle e_2, t_2 \rangle$ is a task execution such that $\mathsf{Exec}(i_2) = e_2$ and $\mathsf{Task}(i_2) = t_2$.

After this initial conversation the senior radiologist decides to show an X-ray of human lungs to the assistant which he finds in a wiki (Wiki, 2006). The X-ray clearly shows symptoms of pneumonia in the lungs together with a textual explanation. Here, the Wikipedia Web site is an actor a_3 which plays the role of knowledge supplier s. Formally, $e_3 = \langle a_3, s \rangle$ is an enactment such that $\mathsf{Player}(e_3) = a_3$ and $\mathsf{Enact}(e_3) = s$. The task *supplying radiology wiki* which is executed by s implies that $i_3 = \langle e_3, t_3 \rangle$ is a task execution such that $\mathsf{Exec}(i_3) = e_3$ and $\mathsf{Task}(i_3) = t_3$.

Characterization between the senior radiologist and Wikipedia can be illustrated by the events in which the senior radiologist searches for usable text and images until the returned results are sufficient to reduce the utilizer's knowledge need. This specific characterization process can be formalized as $\mathsf{MIn}(e_3) \cap \mathsf{MOut}(e_2) \neq \emptyset \vee \mathsf{MIn}(e_2) \cap \mathsf{MOut}(e_3) \neq \emptyset$.

Eventually, the HTTP protocol (as part of the Internet in essence) is an actor a_4 which enacts the role of knowledge transporter p and transports the assets to the utilizer. Hence, the task *transporting radiology wiki* is executed by the transporter. Therefore, $e_4 = \langle a_4, p \rangle$ is an enactment such that $\mathsf{Player}(e_4) = a_4$ and $\mathsf{Enact}(e_4) = p$. Also, $i_4 = \langle e_4, t_4 \rangle$ is a task execution such that $\mathsf{Exec}(i_4) = e_4$ and $\mathsf{Task}(i_4) = t_4$. The exchange of knowledge assets can be expressed by $\mathsf{In}(e_4) \cap \mathsf{Out}(e_3) \neq \emptyset \vee \mathsf{In}(e_1) \cap \mathsf{Out}(e_4) \neq \emptyset$.

A wiki is an example of explicit knowledge. A typical situation of a knowledge market involving implicit knowledge can also be found in the medical domain. Assume that a physician working at the radiology department of a hospital has a relatively good overview of his colleagues' expertise. This physician functions as a broker in the knowledge market. Assume that a colleague has a question about tuberculosis symptoms so that he becomes a potential knowledge utilizer. The local knowledge broker can refer him to another physician who knows more about those symptoms. This colleague will then become a potential knowledge supplier. If the utilizer starts a conversation with the supplier the air which transports the sound of the words functions as a knowledge transporter. If e-mail is used to ask a question then the e-mail system will be the knowledge transporter.

In the next section the discovery and exchange of revealed knowledge is discussed, using a specialization of the knowledge market paradigm.

4 DISCOVERY AND EXCHANGE OF REVEALED KNOWLEDGE

Knowledge discovery of revealed knowledge is complementary to *information retrieval*. The *matching* and transportation of the data which carries this knowledge is a controllable technical issue if every party exactly knows which knowledge is required and offered. Information retrieval is related with the automated or manual search for revealed knowledge which is represented by an explicit characterization. It is certain that the requested data (the carrier of the sought knowledge) is not available if there is no match and the used information retrieval mechanism is properly constructed. However, one may still ponder if one has searched for the right knowledge. In other words: has the query correctly characterized the need for knowledge?

The knowledge discovery paradigm of figure 3 illustrates the matter mentioned above. This paradigm is based on the information discovery paradigm as found in (Proper, 1999). This paradigm is a specialization of the knowledge market paradigm, aimed at the discovery of revealed knowledge. Figure 3 shows a trajectory with the knowledge gap of a utilizer as starting point, or in other words the moment when someone experiences a knowledge gap and the necessity to fill that gap. This leads to a need for knowl-

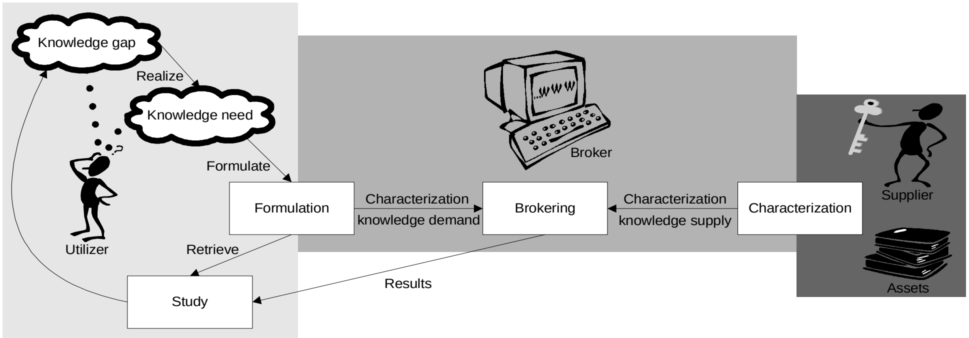

Figure 3: A knowledge discovery paradigm.

edge. It is assumed that a need for knowledge is influenced by what the utilizer already has received from the transporter in terms of assets. As introduced earlier by (Weide and Bommel, 2006), this can be modelled as a function:

$$\mathsf{Need} : \wp(\mathcal{KA}) \times \mathcal{KA} \mapsto [0,1] \tag{10}$$

$\mathsf{Need}(\mathcal{S},a)$ is interpreted as the residual need for knowledge asset a after the set $\mathcal{S}$ has been presented to the utilizer, where $\mathcal{S} \subseteq \mathcal{KA}$. The set $\mathcal{S}$ can be interpreted as the personal knowledge of a utilizer (also called a knowledge profile) during the discovery and exchange of knowledge. No more knowledge is required by the utilizer if his need for knowledge deteriorates, which is denoted by $\mathsf{Need}(\mathcal{S},a) = 0$. To lessen a knowledge need the characterization of a knowledge gap is necessary. This comprises the formulation of a description of this knowledge need in terms of a question (which contains the description but which is not directed to someone), or a query (which communicates the question to a machine). In the knowledge discovery paradigm this is described as the *knowledge demand*. In the ideal case a knowledge demand is formulated by the utilizer and the broker. In a medical context this can be compared to a conversation between an assistant radiologist (who requires specific medical knowledge about a disease) and a senior radiologist. The broker can support the utilizer in formulating the knowledge demand which is a crucial step in knowledge discovery (of revealed knowledge).

The supplier and the deliverable knowledge assets are positioned opposite from the utilizer. Ideally, these assets are described by means of a cooperation between the supplier and the broker in terms of a characterization. The broker can assure that this characterization is also used by the utilizer of this knowledge while formulating their knowledge demand. The brokering activities which are carried out by the broker are in principle not different from an information retrieval process.

In order to have a graphical representation of the definitions discussed in both the knowledge market and discovery paradigms, an object-role model is presented in figure 4. For details on object-role models, see e.g. (Halpin, 2001). Figure 4 depicts a special-

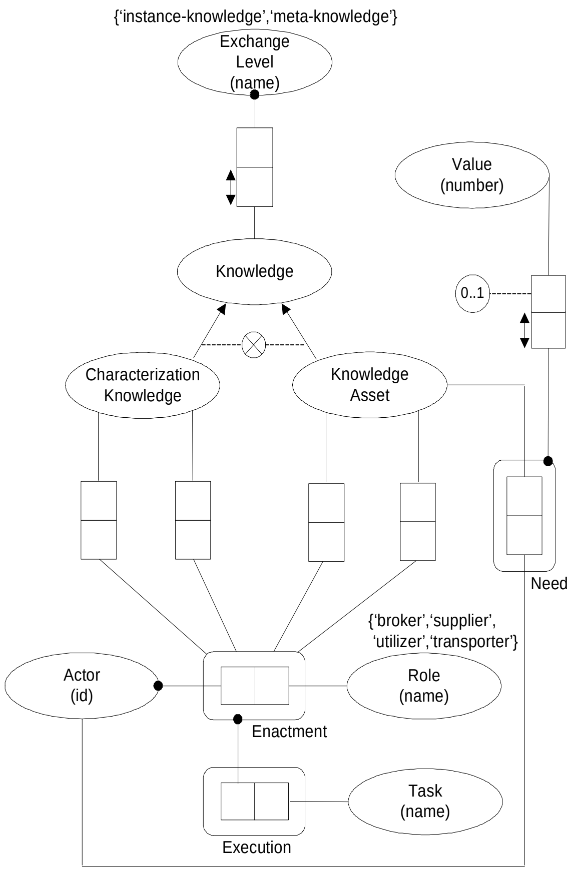

Figure 4: Object-role model of knowledge market and discovery paradigms.

ization relation between the subtypes *Characterization Knowledge* and *Knowledge Asset* and a supertype *Knowledge*. This implies that characterization knowledge and each knowledge asset is, of course, knowledge. For proper specialization, it is required that subtypes are defined in terms of one or more of their supertypes. Such a decision criterion is referred to as a *subtype defining rule*. In figure 4 the subtype defining rule for *Characterization Knowledge* is:

> CharacterizationKnowledge = Knowledge has ExchangeLevel: 'meta-knowledge'

The subtype defining rule for *Knowledge Asset* is:

> KnowledgeAsset = Knowledge has ExchangeLevel: 'instance-knowledge'

The language used in the definitions of the subtype defining rules is described in e.g. (Hofstede et al., 1993).

Thus far we have focussed on a knowledge market paradigm, a scenario of a knowledge market in practice and a knowledge discovery paradigm. In the next section a provisional Web-based application is introduced to deliver support for activities within those paradigms.

5 DEXAR: A WEB-BASED APPLICATION FOR DISCOVERY AND EXCHANGE OF REVEALED KNOWLEDGE

In this section we discuss a way of support for brokering and transporting activities as mentioned in the knowledge market paradigm and partly in the knowledge discovery paradigm, in the form of a provisional Web-based application referred to as DEXAR: *Discovery and eXchange of Revealed knowledge*. Furthermore, we introduce an initial user interface for this application which illustrates interaction between a physician and the application.

The DEXAR application implements a support mechanism so that it keeps track of a utilizer's knowledge profile by collecting a utilizer's knowledge questions together with the knowledge supplied. Feedback of the utilizer to DEXAR creates insight in a utilizer's knowledge need. When a potential utilizer wishes to acquire knowledge he can start to interact with DEXAR . Assume that John Doe is an assistant radiologist who would like to know more about pneumonia while studying at home. John opens his Web browser and starts a conversation with DEXAR , which is shown in figure 5. This is in principle a characterization process in which the application tries to characterize the knowledge demand. After this conversation, the broker application finds an image and text on the Web which might be relevant for John Doe. Figure 6 shows how the application presents part of the results to John. To determine if John still requires additional

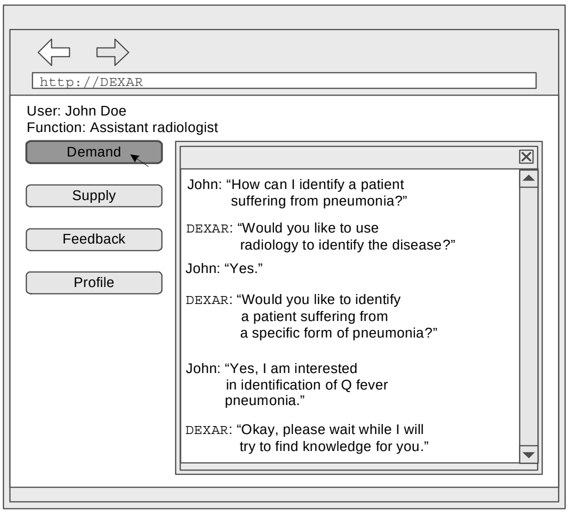

Figure 5: Conversation with DEXAR .

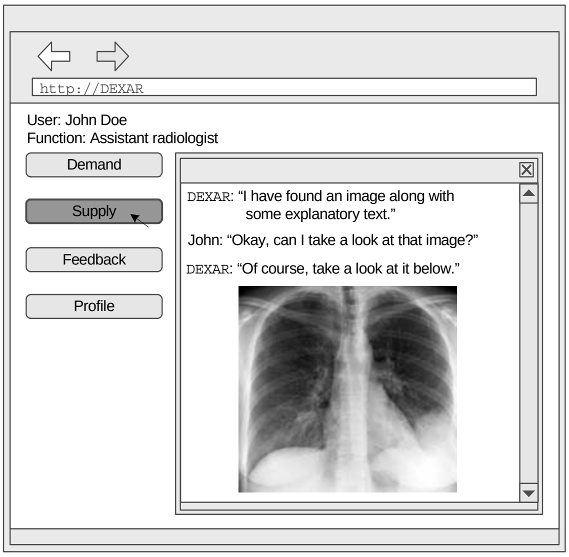

Figure 6: Retrieving knowledge assets with DEXAR .

knowledge, which is the case if $\mathsf{Need}(\mathcal{S},x) > 0$, a feedback screen is shown which is depicted in figure 7. John's final input in the conversation on the screen of figure 7 indicates that $\mathsf{Need}(\mathcal{S},x) = 0$. John Doe can view his knowledge profile which is shown in figure 8. A history is generated showing John Doe's requests for knowledge and the results provided by DEXAR . Underlined words such as the word 'image' represent hyperlinks to the underlying Web pages. So clicking on the word 'image' retrieves that image from the Web. Furthermore, a *lattice* is constructed containing the terms which the application has distilled from conversations with the user. The user may browse through the lattice to learn about previously recorded knowledge and to gain insight in the user's

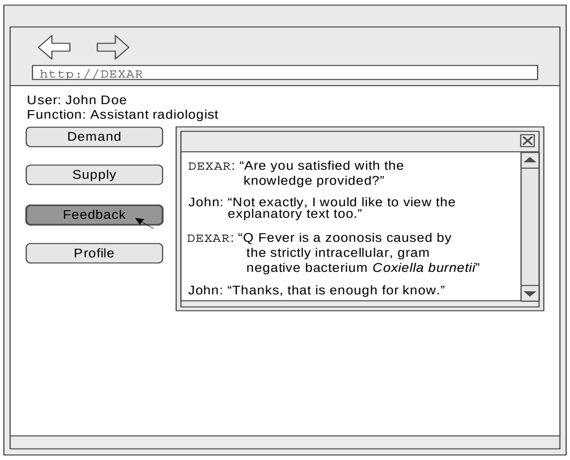

Figure 7: Feedback process in DEXAR .

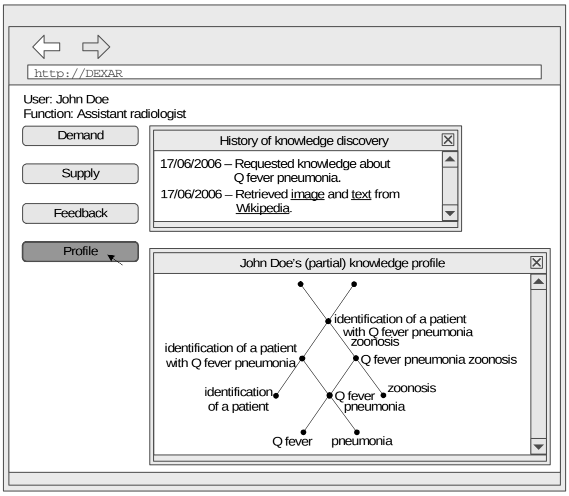

Figure 8: Showing a (partial) knowledge profile.

own profile as a whole. In figure 8, John Doe's knowledge profile is only partially displayed as a lattice. The lattice in DEXAR is constructed by using index expressions as can be found in e.g. (Bruza, 1990). Index expressions have the following syntax:

$$\begin{aligned} \mathsf{IdxExpr} &\rightarrow \mathsf{Term}\{\mathsf{Connector\ IdxExpr}\}^* \\ \mathsf{Term} &\rightarrow \mathsf{String} \\ \mathsf{Connector} &\rightarrow \mathsf{String} \end{aligned}$$

The lattice shown in figure 8 resembles a partially displayed power index expression. A power index expression is the set of all index expressions, including the empty index expression and the most meaningful index expression. An example of an index expression is (identification of a patient) with (Q fever pneumonia). Simply put, (power)index expressions are used by DEXAR as a representation for a knowledge profile.

Searching through a user's own knowledge profile can be implemented by using *Query by Navigation* as is described in e.g. (Grootjen, 2000). At first, the user may provide the application with an index expression (in its shortest form this is a single Term). Once the user is done specifying such a query the application 'knows' which knowledge discovery history (which is coupled to the index expression) can be shown together with hyperlinks to relevant resources on the Web.

Now that theoretical models and a possible way of support for discovery and exchange of revealed knowledge has been discussed, it is appropriate to compare our approach with other approaches in the field. The next section therefore deals with this matter.

6 DISCUSSION

When studying the literature on theory and applications in the area of knowledge discovery and exchange, it is obvious that knowledge discovery, whatever the reason, is often equalled to data mining. Consider for example (Roddick et al., 2003). When compared with the specific distinction made in figure 1, this means that when equalling knowledge discovery with data mining one is focussing on *explicit & concealed knowledge*. We believe that when discussing the status and the type of knowledge one has more comprehension of what can be discovered and exchanged than when solely understanding discovery of knowledge to be data mining. Using the frame of thought as depicted in figure 1 one can probably focus more easily on the discovery of one or more of the knowledge types shown in the four quadrants.

In (Desouza and Awazu, 2003) an *internal knowledge market* is defined as a collection of buyers and sellers who interact to determine the price of a product or a set of products. The main components of an internal knowledge market are: the players (buyers and sellers), rules (governance of interactions), and space (area where buyers and sellers collect). Compared to the knowledge market paradigm discussed in this paper, the *buyers* and *sellers* are in accordance with the utilizer and the supplier roles respectively. However, there are no components that can be compared to the transporter and the broker roles. To be able to focus on the exchange of knowledge within a knowledge market (which is an evident part of the research reported in this paper) analysis of the interaction aimed at delivery and broadcasting of knowledge between buyers and sellers is necessary. An advantage of our model is that those interactions are made clearer by the introduction of additional roles and characterization knowledge and knowledge assets. In our view, a *space* component is not actually a *component* of a knowledge market itself but is dependent of a specific *instantiation* of a knowledge market. I.e. if roles in a knowledge market are executed by actors who are located at a (physical) library then the knowledge market is part of a 'physical' space. If a knowledge market mechanism as proposed in this paper takes part online, other non-physical actors may be involved.

The *rules* shown in the model of (Desouza and Awazu, 2003) however *do* intend to cover exchange mechanisms within a knowledge market. These exchange mechanisms in a market should address what goods will be bought and sold and how they will be paid for. The pricing of knowledge is an important issue to address, but it seems that the approaches of (Desouza and Awazu, 2003) and also (Brydon and Vining, 2006) are more identical to how traditional economic markets function by primarily focussing on price and volume of knowledge. Our models elaborate on the discovery and exchange of knowledge more comprehensively instead because we specifically address those topics and also believe that they are a crucial part of knowledge market mechanisms.

7 CONCLUSION

This paper describes a vision on knowledge discovery and exchange from a conceptual level, illustrated by several reference models. Proceeding from these models a Web-based application illustrates how the reference models can be materialized within the medical domain.

Future research is aimed at the possible application of DEXAR at the radiology department of the Nijmegen University Medical Centre St. Radboud, The Netherlands, EU so that possible positive and negative experiences made with DEXAR can be understood. A goal is to support radiology students fulfilling knowledge intensive tasks. This support is divided in two parts: improving the students' learning process and improving the eventual product of learning, e.g. practical work and exams. Another possible application of DEXAR in the area of information & knowledge systems modelling will be studied. The specific focus in that area is to support modellers in the process of modelling *architecture principles*. This should eventually lead to an improved product i.e. an improved architecture.

REFERENCES

Bruza, P. (1990). Hyperindices: A novel aid for searching in hypermedia. In Rizk, A., Streitz, N., and André, J., editors, *Hypertext: Concepts, Systems and Applications; Proceedings of the European Conference on Hypertext - ECHT 90, number 5 in Cambridge Series on Electronic Publishing*, pages 109–122. INRIA, Paris, France, EU, Cambridge University Press, Cambridge, UK, EU.

Brydon, M. and Vining, A. (2006). Understanding the failure of internal knowledge markets: A framework for diagnosis and improvement. *Information & Management*, 43(8):964–974.

Desouza, K. and Awazu, Y. (2003). Constructing internal knowledge markets: Considerations from mini cases. *International Journal of Information Management*, 23(4):345–353.

Gils, B. v., Bommel, P. v., Proper, H., and Weide, T. v. d. (2006). Quality makes the information market. In *Proceedings of the 14th International Conference on Cooperative Information Systems (CoopIS)*, volume 4275 of *Lecture Notes in Computer Science*, pages 345–359. Springer-Verlag, Berlin, EU.

Grootjen, F. (2000). Employing semantical issues in syntactical navigation. In *Proceedings of the 22nd BCS-IRSG Colloquium on IR Research*, pages 22–33, Sidney Sussex College, Cambridge, UK, EU.

Halpin, T. (2001). *Information Modeling and Relational Databases, from Conceptual Analysis to Logical Design*. Morgan Kaufmann, San Mateo, CA, USA.

Hofstede, A. t., Proper, H., and Weide, T. v. d. (1993). Formal definition of a conceptual language for the description and manipulation of information models. *Information Systems*, 18(7):489–523.

Hoppenbrouwers, S. and Proper, H. (1999). Knowledge discovery: De zoektocht naar verhulde en onthulde kennis. *DB/Magazine*, 10(7):21–25. In Dutch.

ISO/IEC (1994). Information technology - open systems interconnection - basic reference model: The basic model. International Standard ISO/IEC 7498-1:1994(E), Information Technology & Management, Geneva, Switzerland, EU. 2nd Edition.

Liang, T. (1994). The basis entity model: A fundamental theoretical model of information and information processing. *Information Processing & Management*, 30(5):647–661.

Nonaka, I. and Takeuchi, H. (1995). *The Knowledge Creating Company*. Oxford University Press, New York, NY, USA.

Proper, H. (1999). What is information discovery about? *Journal of the American Society for Information Sciences*, 50(9):737–750.

Roddick, J., Fule, P., and Graco, W. (2003). Exploratory medical knowledge discovery: experiences and issues. *ACM SIGKDD Explorations Newsletter*, 5(1):94–99.

Weide, T. P. v. d. and Bommel, P. v. (2006). Measuring the incremental information value of documents. *Information Sciences*, 176(2):91–119.

Wiki (2006). Radiology. Wikipedia, the free encyclopedia. http://en.wikipedia.org/wiki/Radiology.