Implementing a Knowledge Medium in a Multi-centered Clinical Trial

Rolf Grütter
Media and Communications Management Institute, University of St. Gallen, Switzerland
rolf.gruetter@unisg.ch

Katarina Stanoevksa-Slabeva
Media and Communications Management Institute, University of St. Gallen, Switzerland
katarina.stanoevska@unisg.ch

Walter Fierz
Institute for Clinical Microbiology and Immunology, St. Gallen, Switzerland
walter.fierz@gd-ikmi.sg.ch

Abstract

The health care industry is essentially knowledge-based. Despite of this fact, knowledge management and processing techniques are mainly used in form of isolated systems for very specific domains. The basic processes of knowledge generation, distribution, and consumption across domains and locations are not supported by integrated computer systems. Under the growing pressure on quality assurance and cost reduction, concepts and technologies to support the management of knowledge are increasingly gaining the interest of stakeholders. Such a concept – the knowledge medium – is introduced in this paper, and its applicability to the health care domain is demonstrated on the example of a multi-centered clinical trial, i.e. the Swiss HIV Cohort Study.

We define a knowledge medium as a platform for the exchange and management of knowledge within a specific community of agents. It comprises the following components: information objects, i.e. interactive carriers containing externalized knowledge; agents who form a community with common interests and goals; a logical system, which defines the common syntax and semantics of the knowledge managed by the medium.

We conclude that there are great opportunities for knowledge media in the health care domain, and that the development of a knowledge medium must evolve such as to take into account both the processes of knowledge management and the technological convergence.

1. Introduction

Today, corporations are discovering knowledge as their critical asset to successfully compete in emerging global markets. In the realm of this movement, an interesting attempt was made by Arthur [1]. He showed that knowledge-based industries are subject to increasing returns, whereas conventional energy-based industries exhibit diminishing returns.

The health care industry is essentially knowledge-based. Despite of this fact, knowledge management and processing techniques are mainly used in form of isolated (e.g. expert) systems for very specific domains. The basic (repeatable) processes of knowledge generation, distribution, and consumption [17] across domains and locations are not supported by integrated computer systems. Under the growing pressure on quality assurance and cost reduction, concepts and technologies to support the management of knowledge are increasingly gaining the interest of hospital workers, physicians, pharmacists, health insurance companies, and patients.

Knowledge management aims at the timely allocation of externalized knowledge in adequate quality and quantity to the demanding agents. Several technologies, such as organizational memory, have been developed in order to achieve this aim, each of them stressing specific aspects of knowledge management. So far, there is no integrated solution [13].

In this paper we will introduce the concept of knowledge medium, which goes beyond existing solutions for knowledge management and will demonstrate its applicability to the health care domain on the example of a multi-centered clinical trial. The project is a joint effort of the Swiss HIV Cohort Study (SHCS), the Patient-Oriented Medical Information System (POMIS) initiative of Walter Fierz, MD, and the Expert Group Knowledge Media of the Media and Communications Management Institute, University of St. Gallen, Switzerland.

2. Tacit and Explicit Knowledge and the Knowledge Creation Circle

"Understanding how knowledge is developed in a company is a precondition to manage knowledge and
intellectual capital” [18]. Therefore, we will first explain the term knowledge and the knowledge creation circle in accordance with Nonaka [13]. On this basis the concept of knowledge medium will then be motivated.

2.1 Definition and Classification of Knowledge

We define knowledge as the internal state of an agent following the acquisition and processing of information [14]. An agent can be a human being, storing and processing information in his brain, or a virtual machine1.

Human knowledge can be divided into two parts: tacit and explicit knowledge [13]. Tacit knowledge is person dependent and comprises the subjective insights, intuitions and hunches of individuals. Thus, it is hard to formalize and, in consequence, difficult to communicate to others. Tacit knowledge is also deeply rooted in an individual’s commitment to a specific context such as a craft or profession, a particular technology or product market, or the activities of a workgroup or team. In other words, tacit knowledge is deeply ingrained into the context, i.e. the owners view and imagination of the world, and into his experience which is previously acquired knowledge. Therefore, tacit knowledge is a combination of information, context, and experience [9].

Explicit knowledge is coded, i.e. externalized tacit knowledge. Examples of explicit knowledge are documents in enterprises describing projects or other experiences. It consists of three components: a language, information, and a (physical) carrier. The language is used to express and code the knowledge. Information is coded knowledge. Conversely, it is also potential knowledge, which is realized to new tacit knowledge when combined with context and human experience [13]. The carrier physically incorporates coded knowledge and stores, preserves, and transports it through space and over time. During history, several types of carriers have been used as for example paper or computers [13].

Different from tacit knowledge, explicit knowledge is formalized, classified, and available on a carrier, i.e. independent of humans. For this reason, it can be easily communicated, shared, reused, and preserved even if its human creator is not available any more. The usefulness of explicit knowledge depends on its completeness.

The aim of knowledge management is to balance these two types of knowledge and to channel them towards greater innovation and profitability of the organization.

In particular, this means to manage the knowledge creation circle which is explained in detail in the next section.

2.2 The Knowledge Creation Circle

According to Nonaka [13] organizational or community knowledge is created through a continuous dialog between tacit and explicit knowledge (c. figure 1).

![Figure 1. The knowledge creation circle](image)

The starting point is always tacit knowledge, which is created by employees and externalized on an carrier such as paper or computers. Thereby, moving from tacit to explicit knowledge is the articulation of one’s vision of the world. Thus, explicit knowledge represents a specific view on a given world which is related to a given context.

In its explicit form, knowledge can further be combined with other explicit knowledge to new one. At the end of the process is the internalization of explicit knowledge to new tacit knowledge.

From the process point of view the knowledge creation circle is a set of interwoven processes of knowledge creation, externalization, usage, combination, and internalization which are defined around a given kernel of explicit knowledge. These processes are performed by human and artificial agents, in parallel and in different combinations. They result in a new kernel of organizational knowledge being created from a critical mass of new or redefined knowledge.

The knowledge creation circle has, furthermore, two dimensions [13]: an epistemological dimension representing the fluid tacit part of knowledge and an ontological dimension representing the language foundation of knowledge. Languaging, i.e. assigning words to concepts, is the basic means for the externalization of knowledge. Without language,
knowledge can not flow from one person to another within an organization. New experiences or concepts are expressed in new words [18]. These words are then shared in a group or community, might even become part of the organizational terminology, or characterize the communication among different organizations.

Language is usually defined within a context and within a group of persons, i.e. a community. Thus, even though there might exist an organizational terminology, different units might, in addition, use their own, specific vocabulary. In consequence, knowledge management has to consider the particularities of such communities and at the same time provide support for integrated, organizational views.

The knowledge creation circle points out that human carriers of tacit knowledge and artificial carriers of explicit knowledge have to be considered as a single entity and that solutions are required that support a cooperation between them. In the information age this challenge can be anticipated by the new powerful and interactive information carriers resulting from the ongoing convergence of information and telecommunication technologies and traditional media.

3. The Concept of Knowledge Medium

We define a knowledge medium as a platform, which supports the exchange and management of knowledge within a specific community of agents. It comprises the following components (for a detailed and formalized derivation see [10; 14; 15]):

- information objects, i.e. interactive carriers containing externalized knowledge;
- agents who form a community with common interests and goals;
- a logical system, which defines the common syntax and semantics of the knowledge managed by the medium.

The term knowledge medium was introduced by Stefik [17] as an information network with semi-automated services for the generation, distribution, and consumption of knowledge. According to Stefik, the goal of building a knowledge medium is to tie expert systems and communication media together into a greater whole. The concept of knowledge medium presented in this paper extends Stefik’s conception in two directions. First, it does not only consider expert systems but generalizes it for all kinds of knowledge (and information) sources. Second, it explicitly includes the agents who create and use the knowledge medium.

3.1 Information Objects

Information objects are the channels, the interactive carriers containing the externalized knowledge which is exchanged between the communicating agents. This definition requires further precision with respect to the terms involved.

As described above, knowledge – in the narrow sense – must be associated with an agent. Therefore, we avoid to use the crude term when it is tied to a (stateless) carrier. Instead, we use externalized knowledge in those cases.

A carrier includes a physical medium (sound waves in the case of human speech) and applies a language (syntax and semantics) for the logical representation of the contents. In the context of a knowledge medium we rather allocate the syntax and semantics to the logical system. This terminological overlap reflects the fact that the division of a knowledge medium into its components is an artificial one and that the borders of the components are not clear-cut. However, at the meta-level, the meta-syntax and structural semantics (this term will be introduced further on) are an essential part of the carrier. This applies particularly in those cases, where the externalized knowledge is exchanged in document containers.

Interactivity means that the carrier can respond to some extend to inputs and inquiries. This feature distinguishes it from traditional channels like paper documents.

3.2 Community of Agents

As already mentioned, an agent can be a human being, storing and processing information in his brain, or a virtual machine. In the case of human beings, agents form a community with common interests and goals. This community is organized in some way – either explicitly or implicitly.

As a first constituent of their organization, agents take over well-defined roles that demand certain capabilities from them and impose certain obligations and rights. Agents can have artificial representatives in the knowledge medium.

As a second constituent, agents act according to generally accepted protocols. Protocols are all processes necessary for the correct functioning of a knowledge medium. Protocols are defined for roles (and not for agents). They follow predefined rules which regulate and coordinate the interaction of agents in the processes of knowledge generation, distribution and consumption.

3.3 Logical System

As mentioned above, language is the necessary prerequisite for the externalization and exchange of
knowledge. The logical system of a knowledge medium provides the language which is used in order to code knowledge. It comprises the syntax and the semantics. The syntax defines the alphabet, the grammar, and the rules (the axioms) according to which correct sentences must be constructed. The syntax must be implementable on a carrier and applied by the agents of the knowledge medium. The semantics determines the meaning of the used language constructs. It has to be unambiguously defined for the community using the language. The semantics is fixed by reference to real-world objects. Thus, the logical system refers to the same view on the real-world as is shared by the community of agents.

A knowledge medium takes advantage of the ability of the new carriers to support the knowledge creation circle. In order to achieve a symbiosis of artificial and human agents, the language should be understandable for both men and computers. Therefore, it must have a representation in machine-readable form. The machine-readable part is represented within a meta-layer which is added to the contents of the knowledge medium. It enables the structuring and classification of externalized knowledge. This representation can be used by computer programs to retrieve and combine externalized knowledge. In addition, the meta-layer includes meta-information as for example author and address of receiver or sender which are necessary for the control of the communication and – again – for the knowledge creation circle.

The logical system also includes a formal representation of the organizational structure of the knowledge medium allowing for automated reasoning over processes and roles.

In conclusion, the logical system is the binding element between human agents and machines and between tacit and explicit knowledge.

3.4 Relationship between the Components of the Knowledge Medium

The above described components of a knowledge medium are related to each other and form a single entity of human and artificial agents exchanging externalized knowledge within a common semantic space (c. figure 2).

Taken together, they furthermore represent a complete mapping of the knowledge creation circle in the real-world environment onto an artificial world based on information and communication technology and capable to mimic human knowledge creation, distribution, and consumption.

Figure 2. The concept of knowledge medium

The contents of the knowledge medium created using the chosen language is a representation of the world as observed by the community. Any given world can be modeled by different communities using different languages. If a semantic connection between the languages of the resulting knowledge media can be established an exchange of externalized knowledge between them can be facilitated by automated translation from one language into another. In this case, the different knowledge media form a knowledge media net. An example of a knowledge media net for the scientific community is described in [8].

4. Implementing a Knowledge Medium in a Multi-centered Clinical Trial

As mentioned above, information is potential knowledge. The management of information with the help of information and communication technology has a long tradition in organizations. As a result, they usually possess huge quantities of information. Still, from the knowledge management point of view this potential knowledge is mainly unharnessed. The main obstacles hindering an efficient use of information for knowledge creation can be summarized as follows:

1. Externalized knowledge is seen as independent from humans and not as being a part of an integrated knowledge creation circle.
2. Information is not provided in a self-contained form supporting an efficient internalization of externalized knowledge.
3. The information sources are usually designed as isolated islands which are not connected with each other.

In conclusion, one important step towards achieving a knowledge medium is the appropriate combination and connection of different information sources to self-contained information objects. This will be demonstrated on the example of the Swiss HIV Cohort Study (SHCS). The information objects component will serve as a basis where upon further components can be implemented. Before going into more detail, SHCS will briefly be outlined.

4.1 The Swiss HIV Cohort Study

The Swiss HIV Cohort Study (SHCS) involves outpatient clinics of center hospitals (referred to as „Cohort centers”) located in the cities of Basel, Berne, Geneva, Lausanne, Lugano, St. Gallen, and Zurich as well as the Coordination and Data Center in Lausanne. It was initiated in 1987 to collect clinical, laboratory, and epidemiological data on the progression and dissemination of the HIV-infection and to improve the health care services provided to HIV-infected patients [11]. Throughout the study, data are regularly collected at the local Cohort centers and selected sets thereof are anonymously stored in a central database system at the Coordination and Data Center. In recent years, more and more private practitioners joined SHCS to complement the Cohort centers.

Currently, the technical infrastructure supporting the SHCS includes various legacy laboratory systems at the local Cohort centers and a relational database system at the Coordination and Data Center. Anonymous data from/to the Cohort centers are exchanged on paper and also by the use of floppy discs and File Transfer Protocol (FTP) [3]. The knowledge generated from the study is distributed in scientific journals, on the occasion of meetings, and through informal, bilateral contacts.

Actually, it takes several weeks until new examination data are recorded in the central database system and subsequently distributed to the collateral Cohort centers. In consequence, the generation of knowledge based on the study data requires a lot of time. In order to allow for a seamless (electronic) data transfer and database access, thereby speeding up the process of knowledge generation, the prevailing structures and processes of SHCS must be renewed. This is particularly targeted by the development and use of innovative Web-technology.

4.2 Implementing a Knowledge Medium in the Swiss HIV Cohort Study

Based on an overall systems architecture (c. figure 3), a common middleware in terms of an Electronic Study Form (ESF) is being developed as a fist step towards an encompassing computerization. This is done using Extensible Markup Language (XML) [2; 12]. XML is a simplified subset of Standard Generalized Markup Language (SGML, ISO 8879), created particularly for distributed computing on the World Wide Web (WWW). This document-based approach has been chosen in order to deal with the long lifecycle of the information generated throughout the study (XML documents are both human- and machine-readable) and to ensure independence on proprietary software systems in a distributed and heterogeneous environment. Furthermore, it allows for easy integration of private practitioners into the knowledge medium.

Based on an overall systems architecture (c. figure 3), a common middleware in terms of an Electronic Study Form (ESF) is being developed as a fist step towards an encompassing computerization. This is done using Extensible Markup Language (XML) [2; 12]. XML is a simplified subset of Standard Generalized Markup Language (SGML, ISO 8879), created particularly for distributed computing on the World Wide Web (WWW). This document-based approach has been chosen in order to deal with the long lifecycle of the information generated throughout the study (XML documents are both human- and machine-readable) and to ensure independence on proprietary software systems in a distributed and heterogeneous environment. Furthermore, it allows for easy integration of private practitioners into the knowledge medium.

Figure 3. Systems architecture

The XML template is programmed according to the preexisting paper-based study form using an SGML authoring tool requiring some manual adaptations. As to date no Web-browsers are available to fully support XML, a user-interface has to be developed apart. For this purpose, an XML parser is used to parse the exchanged documents. The document components are presented according to a so-called object model (which is merely an Application Programming Interface, API). They are then accessed by JavaScript. The scripts are embedded in
Dynamic HTML, the latter being presented by a commercial browser.

To integrate data from legacy laboratory systems comma-separated text files are generated and embedded in Dynamic HTML using Microsoft’s data-binding technology.

The communication between client and server within the local Intranet employs the Hypertext Transfer Protocol (HTTP). The communication between server and server across Intranets is managed using the File Transfer Protocol (FTP). Although only anonymous data are being transferred, the application of Pretty Good Privacy (PGP), a combination of digital signature and data encryption, is being considered.

5. Relating the Project to the Concept of Knowledge Medium

As mentioned, the ESF middleware refers to the information objects component of our knowledge medium. It serves as a channel to exchange data and externalized knowledge among the participating clinics. Besides contents, XML also provides information on the structure of the document. Thus, an information unit is represented as a tuple consisting of a structure element and a (more or less granular) content element. The structural information is represented in the Document Type Definition (DTD), a declarative part at the beginning of the document (c. figure 4) (alternatively, the DTD can be stored in a separate file and referenced in the document).

```xml
<!DOCTYPE SHCS [ 
<!ELEMENT SHCS (COVERSHEET)> 
<!ELEMENT COVERSHEET (IDENTIFICATION, 
VISIT, (REGISTRATION|END| 
REGULARFOLLOWUP))> 
<!ELEMENT IDENTIFICATION (BIRTHDAY)> 
<!ATTLIST IDENTIFICATION 
NUMBER CDATA #REQUIRED 
GENDER (FEMALE|MALE) #REQUIRED 
HEIGHT CDATA #REQUIRED> 
<!ELEMENT BIRTHDAY (DATE)> 
<!ELEMENT DATE (#PCDATA)> 
<!ATTLIST DATE 
DD CDATA #REQUIRED 
MM CDATA #REQUIRED 
YYYY CDATA #REQUIRED> 
... ]> 
<SHCS> 
<COVERSHEET> 
<IDENTIFICATION NUMBER="14111" 
GENDER="male" HEIGHT="167"> 
<BIRTHDAY> 
<DATE YYYY="1947" MM="10" DD="09"/> 
</BIRTHDAY>
</IDENTIFICATION> 
...
</COVERSHEET> 
</SHCS>
```

Figure 4. XML Document Type Definition and document (excerpts)

The DTD fixes the semantics of the document at the structural level. (In this context the term “structural semantics” refers to the mapping of content information elements onto structural information elements, the latter implementing a tree of hierarchical relationships. For instance, the DTD of figure 4 not only provides the information that „14111“ is an attribute value of the IDENTIFICATION element but also that the IDENTIFICATION element is embedded in the superimposed COVERSHEET element and has the BIRTHDAY element as a child.) This structural semantics allows for the automated execution of symbolic inferences on document components. It can be used by computer systems to manipulate, e.g. to retrieve, the content elements.

Different from the above introduced concept of knowledge medium, the interactivity of the ESF middleware is not yet dissociated from the application. This partly because of technical limitations (as already mentioned, a user-interface has to be developed apart in order to bridge the gap between XML and Dynamic HTML). Another reason therefore is of conceptual nature as the exchanged information objects are not previewed so far to include additional program sequences which will furnish them with self-executing capabilities (according to the object-oriented paradigm).

6. Conclusion and Outlook

The implementation of a knowledge medium in a multi-centered clinical trial is an ambitious initiative. The key point distinguishing a knowledge medium from conventional information systems is its integrated approach, particularly including the self-contained, interactive information objects as well as the (distributed) agents and their organization. A knowledge medium takes advantage from the ongoing convergence of information and telecommunication technologies and traditional media. As a result of this convergence, the use of innovative XML technology to develop the ESF middleware is expected to provide a vast potential with respect to future extensions. We hope to take advantage from third-party initiatives which have been kicked off (and still will be) after the passing of XML as a recommendation by the World Wide Web consortium (W3C) in February 1998. Being equivalent to an
international standard, XML provides the degree of openness and flexibility required to enable the knowledge medium to evolve in the future. As a Web-based technology, it allows for short development cycles and ubiquitous access to contents. In the first step of applying the concept of knowledge medium to the Swiss HIV Cohort Study, we have adopted a bottom-up approach. Based on existing isolated information sources a self-contained information object has been designed and implemented.

A next step will be the installation of local database systems at the Cohort centers to store and manage the XML documents. By the use of object-oriented database systems implementing the tree of structure elements as database schema, it will be possible to navigate through the database and process (structural) semantic queries (besides e.g. full text retrieval). Medical (and administrative) personal will access the contents using different filters reflecting their role within the community and/or the state of the process (i.e. the protocol) of health care delivery. Capable to draw inferences on their contents, these database systems will act themselves as agents and, thus, participate in the processes of knowledge generation, distribution and consumption. In addition, the user-interfaces can be enriched with additional (comfort) features such as to enable not only context sensitivity but also interactive learning from individual user habits.

Once the clinical and laboratory data are available at the local Cohort centers, the knowledge medium can be extended to also include online libraries of expert knowledge and decision support systems. As mentioned in the introduction those systems already exist but are usually isolated applications and not well integrated in the process of health care delivery. The leverage of knowledge medium will support the health care providers in their daily work with HIV-infected patients and ensure that the patients can profit from up-to-date and accurate diagnostic and therapeutic procedures.

Apart from these rather content-oriented aspects of knowledge medium development, the process aspect might become more important in the future. Thus, it will not suffice to exchange single contents (referred to as propositional knowledge) but also to add *procedural* knowledge on how the contents have to be processed when received by an agent (see [16] for a logical architecture of a process-oriented approach for corporate planning). Again, these procedures have to be automated as far as possible.

At the Data and Coordination Center tools with flexible online analytical processing (OLAP) capabilities can be integrated with the central database system besides traditional tools for statistical analysis. These will allow for a dynamic planning of clinical trials thereby reducing coordination time and cost. The technology to develop such tools is one of the core competencies of the Media and Communications Management Institute (see [5; 6] and in the context of clinical trials [4; 7]).

This short outlay of future extension options illustrates two things: First, there are great opportunities for knowledge media, especially in the health care domain. Second, the development of a knowledge medium must proceed in an evolving way taking into account both the processes of knowledge generation, distribution, and consumption and the technological convergence.

7. Acknowledgements

It is particularly due to the enthusiasm of Professor Dr. Beat F. Schmid, chair of the Media and Communications Management Institute, and his passionate elaboration of the presented concept that have inspired us to adopt this framework for our development attempts. Likewise, the assistance of Dr. Ulrike Lechner of the same institute in the formalization of the concept is gratefully acknowledged. The project presented in this paper is generously granted by the Swiss HIV Cohort Study (SHCS).

8. References


