Towards reflective information management

Jari Yli-Hietanen[*]
[*]Digital Media Institute
Tampere University of Technology
jari.yli-hietanen@tut.fi

Samuli Niiranen[*†]
[*]Digital Media Institute
Tampere University of Technology
[†]Visiting at Decision Systems Group
Brigham and Women's Hospital
Harvard Medical School
samuli.niiranen@tut.fi

Abstract

We discuss the evolution of information management and argue that the current paradigm is reaching its limits in terms in ability to provide more utility. Our proposition is that this is due to the nature of current information management tools being built around the paradigm of a document, in the abstract sense of the word, sharing many of their limitations. We discuss reflective information management breaking free from this through a native ability to directly reflect on information.

1 Introduction

Proponents of evolutionary psychology, Steven Pinker being perhaps the most famous example, argue that the mental evolution of primates towards the sentient man was pushed forward by the environments faced by early hunter-gatherers (e.g., see (Pinker, 1999)). The result, Homo Sapiens, is the ultimate hunter-gatherer, a social animal who's mental and physical skills are tuned at co-operation and collaboration at the small scale of hunting parties.

Considering the role of information management in this context, humans are adept at processing and storing information natively when it comes to matters close to the hunter-gatherer life style. Natural spoken language as a tool of collaboration, especially when the collaborating parties share a common ground and history, is an evolutionary adaptation providing an extremely efficient information management tool for the hunter-gatherer (see (Pinker, 2000)). The emergence of complexity, as we currently understand the term, in human societies beyond the ad-hoc organizations of the hunter-gatherer coincided with the growing populations and related urban settings facilitated by the surplus of food created by advancing agriculture. The resulting large, complex societies needed new information management tools to handle the large amounts information related to their management. Written language and its practical manifestation, the written document, emerged as the tool to handle the information overflow created by advancing societies.

The use of documents to convey and store the large mass of information created by advancing agricultural, and later on industrial, societies represented a fundamental break from innate information management. Although documents are written in expressive natural language, they fail to reflect the innate linked nature of information management native to the human mind. When we document something, we almost always lose a part of the original piece of information. Typically some part of the original context, of which we are not necessarily even conscious of at the time of the documentation, is lost in the process. With documentation, the worst case is that relevant information is lost into a sea of unrelated data without meaning and context. The key question arising from these considerations is if continuing on the documentation-centric path is the optimal choice in the contemporary world of the digital computer. Also, although highly expressive, the interpretation of natural language compositions is highly context-sensitive and currently relies on our native disambiguation facilities and a shared common ground with the composer.

What comes after digital documentation in information management? Looking at the contemporary developments of information technology, one key trend is the emergence of linkage. For example, a key power of the World Wide Web is that it pro-

vides a way to dynamically link documents together. The benefits of linkage in the case of digital documents are obvious, especially when the goal is to search and sort them. Linkage is to documents what roads are to cities. Still, linked documents do not provide a panacea for the complex information management needs sprouting around us in business and elsewhere. This is true even if we expand the concept of a digital document to include rigidly coded and structured information in order to solve the problem of ambiguity inherent to natural language as we will discuss later on. It can argued that information management systems built around the concept of a digital document, even a linked and conventionally structured one, have failed to fulfill the promises made in the early age of computing. Could the role of linkage in information be brought to a new level and could this change bring about advantages for the management of information?

Looking at this question from the perspective of business or other organizations provides a number of interesting insights. Organizations targeting complex activities have traditionally been hierarchical with formalized responsibilities and roles. Still, modern management science has raised the question whether this organizational approach is the optimal one in all situations (Wilkinson, 1998). So what is holding back the trend towards decentralization in organizations targeting complex activities? Complex activities involve many pieces of information which current information management tools, based primarily on electronic documents, cannot reflect in a natural way. The result is that a hierarchy is required to gradually transform management decisions down to practical work orders and, on the other hand, to provide reporting of factory floor activities back to the decision makers.

As a practical example, a single work order document, even if available in the context of an enterprise resource planning system, cannot natively reflect on the activity it relates to. Not all the linkages, relevant to the purpose and meaning of the work order, are available. Native means here the ability to obtain the relation between this piece of information and the goals of the organization. Comparing large organizations to single-person ventures it can be argued, that a single-person venture has a relative advantage in information management. Namely, the entrepreneur can natively understand the entire value chain.

2 Reflective information management

A viewpoint we denote ‘reflective information management’ approaches the presented problem by introducing a way to manage information with many dependencies always relevant to complex activities. How can we briefly characterize the approach? There are two key aspects to it. First of all, information is natively handled as an unrestricted linkage of atomic reflections, collected ideally directly from transactions resulting from financial or other concrete decisions made by people and including other information relevant to the activity at hand. The descriptive power of a diverse and large set of direct observations cannot be stressed enough. We argue that the resulting associative power of the linkage can create a foundation for reducing the need for layered, manual information management. Secondly, looking at the operational side of reflective information management, the emphasis is put on human abstraction level, iterative collaboration between humans and information management tools. Simply put, reflective information management tries to replace the middle men reading and writing documents produced in a layered way with the ability to directly obtain the relation between each piece of information and the goals of the activity in question.

How does reflective information management relate to the numerous attempts at intelligent computing and artificial intelligence? In contrast to attempts at a generic intelligent machine, using approaches such as neural networks or expert systems (Russell and Norvig 2003), reflective information management attempts to construct a mechanism supporting people in their information management activities. The focus is moved from standalone, unsupervised intelligent systems towards incrementally and cumulative built tools, which continuously collaborate with users through the efficient and flexible use of abstraction level traversal. Within this collaboration, the key tool is support for reflective communication among collaborating people instead of support for formal documentation.

As stated, the capability of the approach is to enable seamless traversal within different levels of abstraction and different dimensions of an available information mass. The goal is not to automate decision making but to provide people making decisions with context-optimized access to information. The key point is that access to information is provided at human abstraction level in a way that people can concentrate on decision making instead of formatting and interpreting information to and from machine-readable formats. Even more important goal is

the capability to provide the people with the information (e.g., in the form of seemingly simple but highly contextualized messages) currently relevant to their decision making.

Going back to the presented historical review towards the use of documents, in the abstract sense of the word, from natural collaboration, we note that ultimately reflective information management takes the way of organizing human collaboration, also in complex activities, back to its natural form, which is more close to communication among tribesmen than to sterile documentation in hierarchical settings.

3 Unrestricted linkage

Considering the described characteristics of reflective information management, the goal of the approach is human abstraction level information management in settings where the innate human ability is insufficient due to the variety or volume of possibly relevant pieces of information.

Hampton (2003) has discussed how humans use abstraction and context in concept representation. Picking up some key points from his discussion, native human information management involves a flexible abstraction mechanism. Characteristic of this flexibility is that context heavily influences how, for example, a specific heard word is interpreted. Another characteristic he lists is the ability to flexibly add new dimensions to previously learned concepts without hindering the ability to pick the currently most relevant dimensions during a cognitive process. Also, human learning is characterized by the ability to fluidly leverage previously learned concepts and their relations while facing new phenomena.

Considering implications of these characteristics for reflective information management, the key observation is the unbounded nature of human abstraction level information management. This leads into the conclusion that design-while-use is the only possible basic paradigm for reflective information management tools. However, we also limit the scope of the tools by not setting the goal at autonomous or unsupervised operation but rather at the capability of the tools to serve as collaborative agents. The key capability of the tools is thus the ability for abstraction traversal and dimension surfing in an unbounded concept space. This is in contrast to the approach used, for example, in Semantic Web research where concepts are modeled using pre-constructed ontologies.

Expanding on this, conventionally information has had to be rigidly and strictly coded and structured to be machine-processed. This formalization sets a hindrance for information processing as typically the structure for the information has to be designed before processing. This leads at minimum to increased complexity in the case where the operation environment evolves during use. In practice, such information processing systems work satisfactorily only in bounded cases where the boundaries of the case space can be predicted reliably. For these reasons, reflective information management does not rely on structures fixed before use but rather on an evolving, unrestricted linkage of direct observations and other information. Thus reflective information management is suited also for unbounded cases.

We will next describe one possible physical manifestation for the described unrestricted linkage where a set of concept relations with practically unbounded expressive power using natural language vocabulary for concept and relation labels provides the basis for concept modeling. The following is an example of concept relations from health care:

```
12892 "low-density lipoprotein"
(is) "cholesterol"

12893 "LDL" (is) "low-density lipo-
protein"

...

44137 "laboratory result" (for)
"patient" 1146045393687

44138 *44137c (is) "290245-1234"

44139 *44137a (is) "LDL value"

44140 *44139c (is) "89 mg/dL"
```

As seen from this example a concept relation has a syntax similar, for example, to the one used in the W3C RDF standards (see http://www.w3.org/RDF/). Specifically, each relation is a <*unique identification, A, B, C, timestamp*> 5-tuple. Basically, the member B relates the member A and C to each other with an optional timestamp indicating the absolute markup time.

The couplings between the concept relations are either implicit or explicit. For example, an implicit coupling exists between all relations containing a "LDL value" labeled concept through the existence of the similar string. Explicit couplings are defined through labeling A, B and C members with a reference notation. This notation uses relation sequence numbers and A, B or C membership as points of reference. For example, when the member C is la-

beled "*441437c", it indicates a reference to the C member of the relation with identification 441437. It is important to note that the representation does not make any difference between data and metadata.

One mechanism for operation in an unbounded case space is the implicit coupling between natural language labeled concepts. This enables coupling of a new concept relation to already existing concept relations without the need to modify them or to even verify their existence. It should also be noted that since pre-defined ontologies cannot be used with unbounded case spaces, natural language is the only practical source for concept labeling. The unavoidable ambiguity of natural language has conventionally hindered its use in applications involving machine-processing. We propose that instead of formalizing the used language, ambiguity should be solved by firstly utilizing context as available through the associative power of the linkage and the related descriptive power of the direct observations contained in it and secondarily by making a clarifying query to appropriate user. Naturally, the ability to learn from these interactions is a critical facility here.

4 Case: health care

Health care work illustrates especially well the characteristics of a field involved in complex activities where documents have had an important role in information management. Attempts at fully digitalizing the patient record, where the goal has been to improve health work efficiency, have run into major problems, especially if one considers the original goal of increased efficiency (Dorenfest, 2000), (Heeks, 2005), (Wears and Berg, 2005), (Ash et al, 2004), (Kuhn and Giuse, 2001). The introduction of an electronic patient record has perhaps removed the conventional archivist and document courier from the picture but something has been lost at the same time. It can be argued that the electronic record has increased formalization of health care work, due to the use of strictly structured and inflexible electronic patient documents (Berg and Toussaint, 2003). This increased formalization has broken up natural paths of information sharing and work organization. The value of tacit knowledge, which is by definition impossible to formalize (Nonaka and Takeuchi, 1995), and a shared common ground between collaborating health care workers have an essential role in successful and efficient clinical work (Stefanelli, 2001). Of course, it cannot be denied that early information processing systems in health care, as well as in other industries, have been successful at the automation of actuarial work exemplified by repetitive computation of tabular numerical data. To clarify this, processing of documents can be enhanced by digitalization but not all work can be made more efficient by digitalizing documents.

Expanding on the set of concept relations presented in the section 3, we will next present a simple example on how information management operations can be carried out in this context. Specifically, we look at how presenting a simple goal via natural language fragments can provide for contextual information access to information and semantic sensor and action interfacing.

Let us assume that for a patient group laboratory test results for LDL cholesterol are made available in the described representation through a semantic sensor interface mapping a laboratory system's structure and semantics for the LDL values directly to concept relations. Now, a medical professional searching for a specific patient's LDL values initiates a search with the query `"cholesterol 290245"`. Search and traversal algorithms as well as an iterative query process can be applied to provide the professional with the value `"89 mg/dL"`.

Assuming that we describe action interfaces with the same relation mechanism, we can furthermore directly couple the found LDL value and the recipient's address information, for example, to a physical software component capable of sending an e-mail message again with the help of an iterative query process. This way to relate the pieces of information and the actions operating on them enables, at least in principle, a native ability to directly reflect on information.

5 Discussion

The current discussion represents merely a starting point and a vision for more detailed work in the area of human abstraction level information management. Important open research questions include, but are not limited to, performance issues related to the utilization an unrestricted linkage and the repeatability of operations carried out in the context of the described paradigm.

References

Steven Pinker. *How the Mind Works*. W.W. Norton & Company, New York, 1999.

Steven Pinker. *The Language Instinct: How the Mind Creates Language.* Harper Perennial Modern Classics, New York, 2000.

Adrian Wilkinson. Empowerment: theory and practice. *Personnel Review*. 27(1): 40-56, 1998.

Stuart Russell, Peter Norvig. *Artifical Intelligence: A Modern Approach.* Prentice Hall, 2nd edition, Upper Saddle River, 2003.

James A. Hampton. Abstraction and context in concept representation. Philosophical Transactions of the Royal Society of London B. 358: 1251-1259, 2003.

Sheldon Dorenfest. The decade of the '90s. Poor use of IT investment contributes to the growing healthcare crisis. *Health Care Informatics*.17(8):64–7, 2000.

Richard Heeks. Health information systems: Failure, success and improvisation. *International Journal of Medical Informatics*. Aug 19, 2005.

Robert L. Wears, Marc Berg. Computer technology and clinical work: still waiting for Godot. *Journal of American Medical Association*. 293(10): 1261-3, 2005.

Joan S. Ash, Marc Berg, Enrico Coiera. Some unintended consequences of information technology in health care: the nature of patient care information system-related errors. *Journal of American Medical Informatics Association*. Mar-Apr;11(2):104-12, 2004.

Klaus A. Kuhn, Dario A. Giuse. From hospital information systems to health information systems. Problems, challenges, perspectives. *Methods of Information Medicine*. 40(4):275-87, 2001.

Marc Berg, Pieter Toussaint. The mantra of modeling and the forgotten powers of paper: A sociotechnical view on the development of process-oriented ICT in health care. *International Journal of Medical Informatics*. Mar;69(2-3):223-34, 2003.

Ikujiro Nonaka, Hirotaka Takeuchi. *The Knowledge-Creating Company. How Japanese Companies Create the Dynamics of Innovation*. Oxford University Press, New York, 1995.

Mario Stefanelli. The socio-organizational age of artificial intelligence in medicine. *Artificial Intelligence in Medicine*. 23(1): 25-47, 2001.