logistyka informacji,
systemy przetwarzania wiedzy,
procesy przetwarzania informacji, ontologie

Bogdan D. CZEJDO[1]
Mikołaj BASZUN[2]

KNOWLEDGE PROCESSING BASED ON INFORMATION LOGISTICS MODELS

The general requirements of information logistics are to provide the right information 'product', in the right format, at the right place, at the right time, and for the right people. In this paper we discuss two information logistics models: (a) hierarchical collaborative knowledge processing and (b) network collaborative knowledge processing. Ontology can be of crucial importance for information processing since most of the information is "unstructured". The iterative process of knowledge processing can be supported by an appropriate software system.

SYSTEM PRZETWARZANIA WIEDZY OPARTY NA MODELU LOGISTYKI INFORMACJI

Generalne wymogi dla logistyki informacji obejmują zapewnienie dostarczenia właściwego „produktu" informacji, we właściwym formacie, w odpowiednim czasie, i do właściwego odbiorcy. W niniejszej pracy przedyskutowane różne modele logistyki informacji. Przetwarzanie informacji jest dyskutowane dla hierarchicznego i sieciowego modelu kooperacji. Ontologia może mieć kluczowe znaczenie w procesie iteracyjnego przetwarzania wiedzy, ze względu na to, że większość informacji nie ma ustalonej struktury. Oparty na ontologii proces iteracyjnego dostosowywania informacji może być wspomagany przez odpowiedni system oprogramowania.

1. INTRODUCTION

Collaborative Knowledge Processing is an active research area with many theoretical and practical results [1], [2], [3], [4], [5], [6], [7]. One of the best known results of collaborative knowledge processing is the online encyclopaedia called Wikipedia. The model of collaborative knowledge processing in Wikipedia is based on two principles: (a) any member of the online community [8] can participate in collaborative work on the

[1] Department of Mathematics and Computer Science, Fayetteville State University, Fayetteville, NC 28301, USA, bczejdo@uncfsu.edu
[2] Mikolaj Baszun, Warsaw University of Technology, Department of Electronics and Information Technology, Nowowiejska 15/19, 00-665 Warsaw, Tel. + 48 22 2347906, mbaszun@elka.pw.edu.pl

encyclopaedia components creation and (b) the members of the online community decide by themselves what new knowledge component could be included. This model is therefore highly dynamic. The software is used to assist participants to create and edit the knowledge components. The model of collaborative knowledge processing in Wikipedia was analyzed in [1] by using agent simulation techniques. The knowledge components (articles) in Wikipedia can differ in content growth rate, number and frequency of updates, and frequency of vandalisms [9]. Generally, the visualizations of dynamic models are very insightful as described in [10], [11].

In this paper we focus on models of collaborative knowledge processing based more controlled environment. The controlled environment is typical for the most of the companies. The companies have well defined organizational/task hierarchy that imposes constraints on collaborative knowledge processing. Employees participating in knowledge processing are referred as knowledge workers [12]. Their companies e.g. IBM are looking for solutions that will have direct impact on the daily performance such workers [13], [14].

There are some promising attempts to apply analogy of controlled knowledge processing to Physical Logistics models as was described in [12], [13]. The Information Logistics requirements are to provide the right information 'product', in the right format, at the right place, at the right time, and for the right people. The storage space and the distance to sent information is of lesser importance. The biggest bottlenecks are related with information transformation and efficiency of knowledge workers. In addition to bottlenecks we have to deal with new challenges such as information overload [12]

One of the biggest challenges is to support vast knowledge "warehouses". Such information repositories contain knowledge in the form of emails, presentations, texts, videos, web pages, illustrations, drawings, photos, images etc. Most of such knowledge is referred to as unstructured data/information since it does not fit neatly in the database and can hardly be managed in spreadsheets, or like systems [12]. But unstructured information is what largely drives time to market and sales and it is necessary for right decisions [12].

The important task is to refine knowledge processing models so that they can be used to improve efficiency of knowledge workers processing unstructured information. There are attempts to classify knowledge into knowledge components in order to describe the cooperation more precisely [12]. In our paper we propose to create a more precise knowledge description based on ontology.

The ontology system, if implemented, should result in some kind of knowledge structure represented by objects, their properties and relationships, and rules. Nature of each of these components called ontology components can be quite different: textual or numerical describing people, material, geometrical shapes, electrical, architectural, etc.

The detailed analysis of the modelled world including some detailed components is crucial for creation of an ontology system. Such analysis should lead to the creation of an initial model of the ontology and help later in ontology transformation.

With ontology, Bottleneck Analysis can be used for information elements and information chains – similarly to supply chains for physical goods. Currently we cannot yet create the computing technology to convert completely unstructured text into complete knowledge structures. We can only use incomplete knowledge structures in the form of ontology components to improve information logistics. The unstructured text can have a "shadow" ontology. This way the ontology can support information processing. There are many information logistics processes that can be improved based on ontology. One

approach considered in this paper is to replace unstructured requests by ontology based requests. Also, ontology based answers could be required or at least providing link between unstructured answers by an ontology. The well defined ontology based information processing can be a foundation for software to provide support for knowledge workers.

The paper is organized as follows. In the next Section 2, the models for cooperation of knowledge workers are discussed in general. In Section 3, the hierarchical cooperation model is presented. In Section 4, the network cooperation model is presented. Summary is given in Section 4.

2. MODELS FOR COOPERATION OF KNOWLEDGE WORKERS

The Knowledge Worker Cooperation models can be classified into static models and dynamic models. The dynamic cooperation models generally assume that the cooperation during knowledge processing cannot and should not be completely pre-specified. There are two aspects of a model that can be dynamic: the choice of knowledge workers and the choice of knowledge content (scope). The fully dynamic models, assume that all knowledge workers willing to participate can participate, and the scope of the knowledge is not pre-determined. The participants decide themselves if they can contribute to the process collaborative work on knowledge processing and what knowledge content they want to work on. Typically software is used that allow participants to easily join the process and contribute to the knowledge recording. The success of Wikipedia was to the big extend possible because the use of fully dynamic model for knowledge recording and processing.

There are some other applications using the fully dynamic model, but in the most companies some constraints are imposed. The typical constraint is related to the scope of the processed knowledge. The companies in most cases require some specific knowledge results (scope), e.g. recommended pay or specific explanation for the public relations action. It means that this aspect of the fully dynamic model is almost never of any serious value. To be more specific, temporary generation of possibly unneeded knowledge may be allowed, but it still needs to serve some way the final goal at the end.

The second aspect of a dynamic model is related with the choice of participants. In such a model, the participants (knowledge workers) can participate at will. Let us consider a situation requiring solutions for a non-typical problem. Let us assume that for some reason e.g. speed, the cooperative knowledge processing is based on email. When the decision maker requests some information he/she sends e-mail to several knowledge workers. The email is copied to several other people. Some of “other people” might need to invite other colleagues to help the process. As was described in [12] an e-mail storm can emerge with dynamically expanding the network of cooperating Knowledge Workers.

The static models assume fixed procedures of knowledge processing determined typically by an organizational/task structure in the company. The static models can be classified further into hierarchical models and network models. In the hierarchical models the hierarchical structure of the company can be mapped into a hierarchy of knowledge flow. In the pure hierarchical model the requests are coming from the top and the processed knowledge flows back to the top.

3. KNOWLEDGE PROCESSING IN A STATIC HIERARCHICAL COOPERATION MODEL

An Knowledge Processing System based on hierarchical cooperation model is shown in Figure 1. The system includes Main Knowledge Worker, other Knowledge Workers, Background Knowledge Repository, and Ontology.

The Main Knowledge Worker and other Knowledge Workers cooperate within pre-determined hierarchy typically corresponding to an organizational or task force structure in the company. In our example we assume a Main Knowledge Worker responsible for a decision requiring knowledge to be provided by other Knowledge Worker according to the hierarchical structure as shown in Figure 1. This is a pure hierarchical model since the requests are coming from the top and the processed knowledge flows to the top.

The knowledge requests usually refer to some repository of knowledge. The knowledge repository can be in the form of explicit documents or references to the documents. The documents can be in the form of emails, presentations, texts, videos, web pages, illustrations, drawings, photos, images etc. The Knowledge Repository is an important component of the system since it creates a common framework of reference to all participating Knowledge Workers. It not only provides useful information but also allows knowledge workers to store a new knowledge and make the new knowledge available to other workers.

The challenge is how to refine knowledge processing models so that they can be used to improve efficiency of knowledge processing. The knowledge processing system described in this paper and reflected in Figure 1 includes Ontology as an important component of the system. Ontology can provide some kind of structure for the knowledge repository identifying knowledge objects, their properties and relationships, and rules. With ontology, Bottleneck Analysis [12] can be used for information elements and information chains – similarly to supply chains for physical goods.

Let us consider in some detain a knowledge processing in Figure 1. The Main Knowledge Worker responsible for a decision sends knowledge requests to Knowledge Workers 3 and 5. The knowledge requests can be sent in the form of the structured ontology requests or unstructured requests. First of all, Knowledge Workers 3 and 5 will be responsible to converting unstructured requests into structured ontology requests if necessary.

The knowledge requests are often accompanied by either reference to some knowledge in the repository or by a new structured or unstructured knowledge. Again, Knowledge Workers 3 and 5 will be responsible to check if the new knowledge is in the repository and/or ontology and update both of them if necessary. In the last phase, the knowledge processing leading to the requested knowledge are performed based again on ontology and its relationships with available unstructured documents in the repository. Even the incomplete ontology knowledge structures can be of big assistance in performing such task. The unstructured documents can have “shadow” ontology supporting knowledge processing. Also, ontology based answers could help in knowledge integration performed by knowledge workers in the upper part of cooperation hierarchy. The ontology based knowledge processing can be supported by software to inform about links between background repository and ontology components.

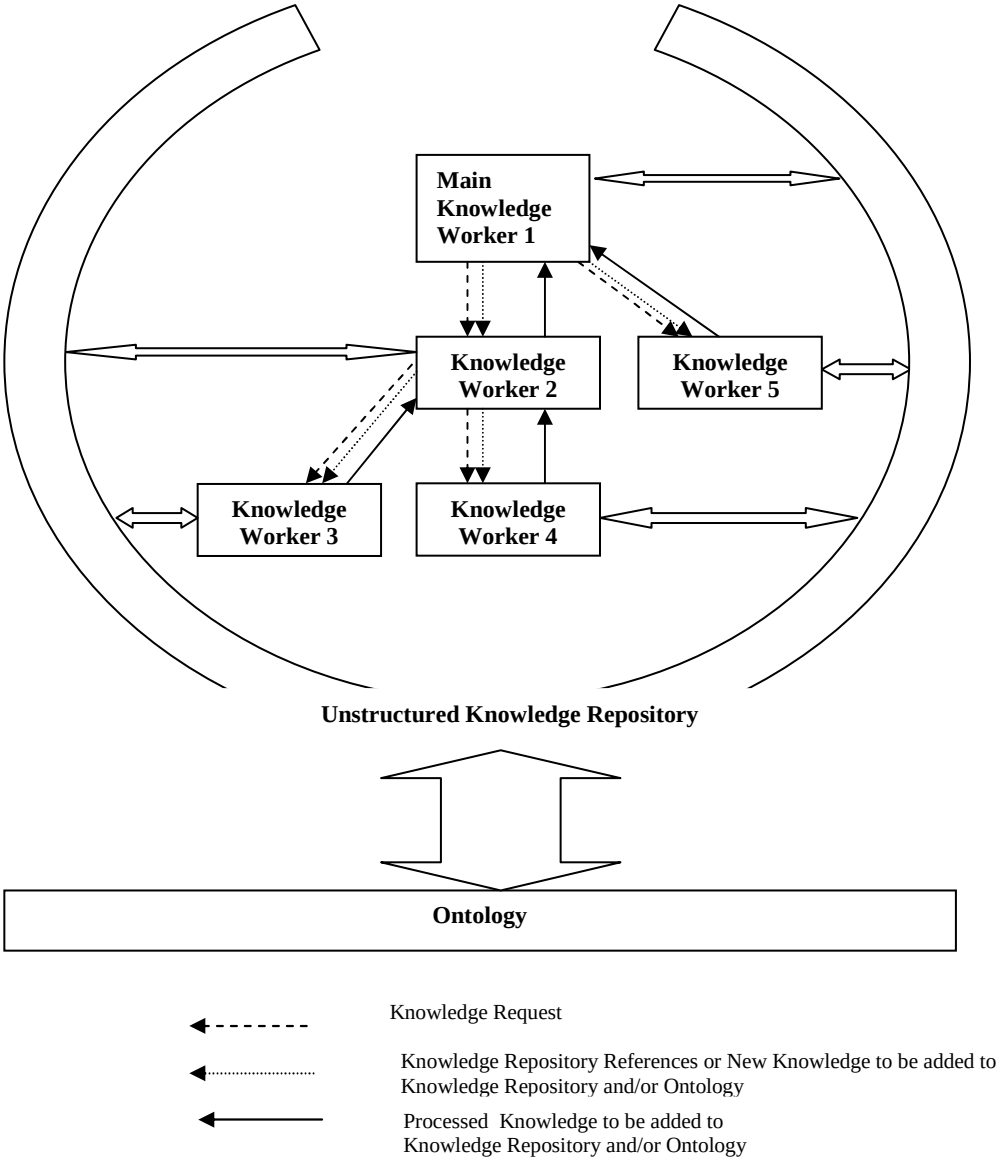

Figure 1. A hierarchical model for cooperation of Knowledge Workers

A pure hierarchical model is not always practical. The model can be made more flexible by allowing cooperation between knowledge co-workers as shown in Figure 2. This will lead us into network cooperation models as described in the next section.

3. KNOWLEDGE PROCESSING IN A STATIC NETWORK COOPERATION MODEL

The network cooperation model allows a Knowledge Worker to request a assistance from any other Knowledge Worker as shown in Figure 2. The biggest advantage is that the knowledge processing task can be matched better with the expertise of another worker. One of the important roles of the ontology is to link workers expertise with structured and unstructured information.

The network cooperation model, introduce new opportunities but also new challenges. Such cooperation model can have, and usually has, cycles (iterations) that can cause divergence of knowledge processing. The convergence of knowledge "solutions" are of a crucial importance. The ontology can provide support to control this problem by providing some measures of distance to the knowledge goal.

Let us consider in some detain a knowledge processing in Figure 2. Let us again assume that, the Main Knowledge Worker responsible for a decision sends knowledge requests to Knowledge Workers 3 and 5. As before, Knowledge Workers 3 and 5 will be responsible to converting unstructured requests into structured ontology requests if necessary. The further processing is different though. Knowledge Worker 3 after some knowledge processing sends his knowledge requests to Knowledge Workers 2 with the information that the result should be send back to Main Knowledge Worker. Knowledge Worker 3 because of some reason e.g. is not available, sends the same knowledge requests to Knowledge Workers 4 again with the information that the result should be send back to Main Knowledge Worker.

4. SUMMARY

The paper discussed methodology to support knowledge processing based on Information Logistics models. We classified Information Logistics models and described architecture of a system to help knowledge workers to "converge" the knowledge to the state in which it can be used to make decisions. The knowledge repository contains unstructured information. A "shadowing" ontology is used to provide a structured "skeleton" to unstructured information. The presented knowledge processing system supports iterative processing. The knowledge workers could be assisted by software supporting different cooperation models.

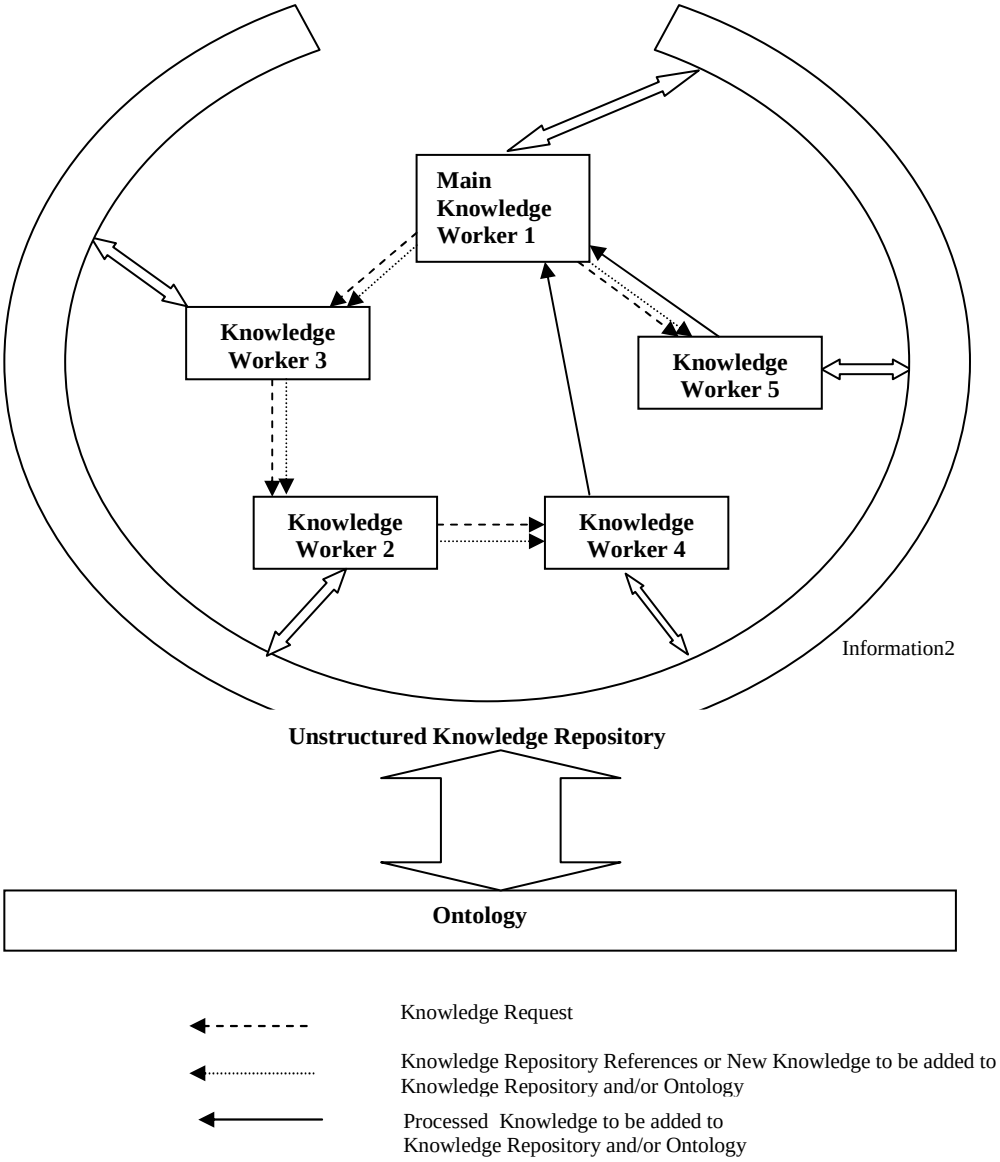

Figure 2. A Network model for cooperation of Knowledge Workers

REFERENCES

[1] Jinsheng Xu, Levent Yilmaz, Jinghua Zhang, "Agent simulation of collaborative knowledge processing in Wikipedia", Proceedings of the 2008 Spring Simulation Multiconference SpringSim 2008, Agent-directed simulation symposium (ADSS'08), Ottawa, Canada, pp.19-25.
[2] Lim, E., Vuong, B., Lauw, H.W., Sun A., "Measuring Qualities of Articles Contributed by Online Communities", Proceedings of International Conference on Web Intelligence, 2006.
[3] Hargadon A. and Bechky B. (2006), "When Collections of Creatives Become Creative Collectives: A Field Study of Problem Solving at Work," Organization Science, vol 17, pp. 177-192.
[4] Diehl, M., and Stroebe, W. (1991). "Productivity loss in idea-generating groups: Tracking down the blocking effect", Journal of Personality and Social Psychology, 61, 392-403.
[5] Gallupe, R. B., Bastianutti, L. M., and Cooper, W. H. (1991). "Unlocking Brainstorming", Journal of Applied Psychology, 76, 137-142.
[6] Dennis, A. R. and Valacich, J. S. (1993), "Computer brainstorms: More heads are better than one", Journal of Applied Psychology, 78, 531-537.
[7] Von Hippel E. (2005), Democratizing Innovation, MIT Press. Cambridge, Massachusetts
[8] Online community, http://en.wikipedia.org/wiki/Virtual_community
[9] Zeng, H., Alhossaini, M., Fikes, R., McGuinness, D., "Mining Revision Histories to Assess Trustworthiness of Article Fragments", Proceedings of International Conference on Collaborative Computing: Networking, Applications and Work Sharing, 2006.
[10] Viégas, F.B., Wattenberg, W., Dave K., "Studying Cooperation and Conflict between Authors with history flow Visualization", Proceedings of Conference on Human Factors in Computing Systems (CHI) 2004.
[11] History Flow Visualization, http://www.research.ibm.com/visual/projects/history_flow/.
[12] Jan Willems, "Information Logistics: From Having to Using", CIO Strategic Event, Nice, September 10-11, 2007.
[13] Jan Willems, "Information Logistics: gaining performance – smiling customers", White Paper NRG Group Benelux, January 2007.
[14] Eduard Feigenbaum, Pamela McCorduck, H. Penny Ni, "The Rise of the Expert Company", Time Books, New York 1988, ISBN 0-679-72518-0.

This work was partly supported by:
- Belk Foundation, North Carolina
- Funds on Science in 2007 – 2010 as Ordered Research Project of Polish Ministry of Science and Higher Educations.