BUDAPESTI CORVINUS EGYETEM

Ph.D. program of Management and Business Administration

COLLECTION OF THESES
for

Barna Kovács

Reducing information overload in organizational workflow systems

Ph.D. dissertation

Supervisor:

dr. András Gábor
associate professor

Budapest, 2009

BCE, Department of information systems

COLLECTION OF THESES
for

Barna Kovács

Reducing information overload in organizational workflow systems

Ph.D. dissertation

Supervisor:
dr. András Gábor
associate professor

Contents

I Antecendents and subject of research

One of the greatest challenge these days is to be able to handle the vast amount of information that is produced on a daily basis. This issue arises on every areas of life, in the case of individuals as well as corporate life and public administration. It should be noted however, that computers and information systems applied in these fields—although capable of handling huge piles of information by themselves—do not solve the problem, moreover they contribute to the escalation of the phenomenon. Finally, the bottleneck is always a human being.

Information processing capacity of human beings is limited, moreover it can also be stated to be moderate. Human thinking, the human brain does not shine in this type of information processing; it is much more talented in the thinking in nexus, network of entity relationships (as presented by Vas 2007, p.117). Accordingly, the challenge is not to empower humans with computer-like information processing skills, but in the other way around, enabling computers to interpret and understand the processed data and information, providing this way such kind of information to its users that are more suitable to solve his or her problem. This way, computers could take over grater share of the user's problem solving activities and find answers based on the stored data and information, or at least revealing the relevant sphere of data and information.

I.1 Justification of the research topic

The fast development in the field of economics, technology and informatics requires the ability of fast adaptation of individuals and organization, being either public administration institutions, corporations or even citizens themselves. Human workforce, however, is not capable any more to process the amount of information that is provided by technology.

Information overload can have a lot of consequences regarding organizations. It jeopardizes the organizations' capability to adapt to economic or market changes, endangers competitive edge. In the same time, it can also cause overloading of employees. Moreover, not only the amount, but the complexity of information is rising that provides even more challenge to organizations.

When information handling processes of an organization are investigated, it can be observed that information systems take part in the processes but mainly as storage facilities for information. In the majority of the cases, role of these systems is to store pieces of information and offer some facilities to retrieve them. In most of the organizations, the problem is that retrieving

information is extremely inefficient using these tools. Search facilities used currently are able to deliver relevant information only in a small percentage of the cases, which means that information retrieval takes really a lot of time and cost for the organization.

It is also an important observation that the majority of data and information stored in the information systems of an organizations is created inside the organization in most of the cases, only a minor part originates in the outside world. Considering a corporation as an example, contracts, studies, offers are created and stored in the organizational information system. It is a common practice to create and store these documents again when needed in a different case, instead of reusing existing documents. Without employing appropriate search methods, employees are not able to find the older documents, they are therefore unable to reuse them.

It is important therefore that research and development must turn to the development of innovative and modern technologies that enable individuals and businesses to find relevant information. Especially in the field of public administration, the effectiveness of which needs to be improved significantly, as heading towards the concept of servicing state.

This direction of research means currently revealing the application possibilities of semantic technologies that will be reinforced by the research of this dissertation as well.

I.2 Research objectives

The research aims to grasp and decrease (possibly minimize) information overload in an organizational environment. This aim is approached from the information system side, since—according to the literature review of the dissertation—information systems provide fairly large contribution to the overload situation. For reaching this goal, a number of improvements are needed considering the current state of information systems. This dissertation analyses an approach that provides promising results.

Roots of the problem lie in the need of human information processing. It is not a problem, when an excess of information is present; the difficulty is caused when a large part of this information needs to be processed by a human being. This situation leads to the concept of information overload. This term can be interpreted on the individual level and on the organizational level as well. In this work, the organizational view has been selected for analysis.

Information overload is not easy to grasp as a phenomenon. One of the fundamental approaches—which was used in this work as well—is to analyse how information can reach workers in the organization. According to this approach, the act of looking for information (information retrieval, searching,

or information pull in other words) and delivering of information (information push) can be distincted. It has to be noted that both approaches can cause information overload by their own when applied separately. A favorable state seems to be a combination of information push and pull. Therefore, the aim of this research is to offer a methodology for setting an optimal rate of information push and pull in order to minimize information overload on the organizational level. According to the theoretical foundations of the dissertation, starting point of the research can be determined as follows:

1. Key element of decreasing information overload is setting of the balance between information push and pull. Extremities are not favorable, instead a combination in the middle should be defined and an appropriate ratio has to be determined.

2. Information has to be structured in order to offer better support to the user. Applying ontologies as description framework for structured information seem to be a proper approach.

3. Intelligent agents seem to serve as an appropriate framework for making this structured information push and pull operate.

II Applied methodologies

Determining the appropriate methodology to the above described problem requires a complex approach. On one hand, the formulation of a methodology is needed regarding the optimization of the information push and pull approaches. On the other hand, a system has to be constructed that enables the realization of the aims of the methodology. First, the applicability of quantitative and qualitative methodologies are investigated, then more details of the selected methodology are presented.

Quantitative techniques are mainly used in researches where large amount of data is used and analysed by applying statistical methods (Füstös et al. 1986). In the case of this research scenario, obtaining appropriate numeric data is not feasible, since the research aims the construction of methodology and a new system, therefore numeric data about the usage of the system are missing during the research.

By using qualitative methods however, the researcher has the opportunity to understand the subject more deeply, actively listen to the researched individuals, and grasp their problem and the problem context. Moreover, not only the pre-defined questions can be answered by the analysis, but new relationships can be explored that were not known in the beginning of the research (Oakley 1999). However, using qualitative methods does not mean that numeric data cannot be used. It only means that the primary goal of the research is to explore and understand context and relationships more deeply, instead of gathering data (Blaxter et al. 2006). It can be seen this way that qualitative methodology is appropriate for the research.

The next question is the selection of the logic for the justification of the aims of the research. As Babbie (1996) suggests, deductive research methodology that formulates and tests hypotheses can be applied successfully, when the field of research has a solid, well-developed theoretical background. In a less-defined area, the researcher can take more risk by formulating hypotheses. In this case, it is advisable to develop a theory that can be employed as a basis of future hypothesis testing.

Inductive methodology, on the other hand, aims at exploring new problem areas and constructing new theories. Applying this method, hypotheses cannot be formulated in the classical sense, relationships has to be built instead, according to the investigation of the problem area aiming to formulate conditional conclusions (Babbie 1996).

The scope of this research lies inbetween the two approaches. It is built on existing theoretical foundations, moreover it uses theories from multiple disciplines; however it also aims to develop a new theory based on them in order to develop the methodology of minimizing information overload. Since

the application of already existing theoretical foundations is more emphasized in the research, it requires deductive approach. In other words, the research justifies existing theories by applying them in another discipline.

II.1 Research strategies

Among the research strategies distinguished by Yin (1994, p.5-9), a combination of exploratory and explanatory strategies can be applied in this research. *Explanatory* research involves answering the "how" and "why" questions, and tries to find an explanation of the phenomena being the subject of the research. For this type of research, using case studies is suggested by the author. (Yin 1994, p.9)

Exploratory research strategies on the other hand answer "what" question —in the sense of "What can be learned from. . . "—focusing on the development of hypotheses and propositions for further inquiry.

Additionally, Eisenhardt (1989) emphasizes that case studies can also be applied in an explorative manner: providing illustration for a theory, construction of a theory, and testing or extending a theory. Research can use multiple case studies or build on only one. Using a single case study can be used to justify or falsify a theory or demonstrating special or irregular cases. As Tellis (1997) summarizes, both single and multiple case studies are acceptable as far as they meet research objectives.

It has to be noted, that Babbie (1996) also delineates drawbacks of case studies: they are hard to repeat, findings are not so unequivocal, and the trustiness of cases is a matter of discussion. As Yin (1994) points out, generalization of results can made only to theory, not to populations.

The nature of this research requires the combination of these two approaches and suggests the use of the case study methodology. As the goal of the research is to investigate the realization of theoretical findings in a specialized pilot case, the application and analysis of a single case study seems to be an appropriate methodology. The outcome of the research will be a single solution that is only one of the almost endless continuum of possible solutions. During the realization, some features of exploratory research have to be applied in the way meant by Eisenhardt (1989), and also supported by Hart (1998). According to the classification of Stake (1995), the type of this research is intrinsic since the researcher has an interest in the case.

III Achievements of the research

The aim of this research is to offer a methodology for setting an optimal rate of information push and pull in order to decrease information overload in the organizational. The research can be considered successful, since its aims have been reached. A methodology has been developed to determine the optimal ratio of information push and pull, moreover it was shown that this ratio provides the minimum of information overload in the given workflow process. It was also shown that realizing information push requires structured information. Structuring information can be realized by using many tools available. It is shown in the dissertation that ontologies provide the most advanced framework for structuring information regardless of its form, moreover employing ontologies also enables the integration of information, which can serve as the basis of information push services encompassing all kinds of information in the organization. In order to justify this statement, a system has been realized that integrates information stored in various types of systems—e.g. content management, groupware and workflow management systems—while offering domain-specific description as well, which is used for supporting highly knowledge intensive tasks, realizing this way an information push service. It is shown that this way of information push can be customized by changing the granularity of the ontology, being independent this way from the other implementation details.

Major achievements of the dissertation are detailed in the following sections.

III.1 Developing a methodology to determine the optimal ratio of push and pull in a workflow system

Aim of the development of this methodology is to effectively decrease information overload—as it is formulated in the main hypothesis of the dissertation. This formulation is needed in order to determine the frames, in which the decrease of information overload can be accepted. For this purpose, the definition of effectiveness has been clarified and adapted to information systems.

III.1.1 Effectiveness

Determining effectiveness was one of the main challenges of the research. The conclusion of the analysis was that effectiveness itself is the "degree to which objectives are achieved and the extent to which targeted problems are resolved" (BusinessDictionary.com 2009). This implies that effectiveness

cannot be interpreted without goals to achieve. During the analysis, the conclusion was drawn that having organization, context and situation are necessary conditions of determining objectives of an organization.

The various approaches can be summarized that establishing goals require an analysis, in which the characteristics of the organization are compared to the business context, e.g. the environment. Basis of this analysis is always a snapshot of the current situation of the organization. In the context of strategic planning, goals can be either the purpose, goals or objectives of the organization. When objectives are determined, effectiveness of the organization can be measured.

However, since subjects of the research are information systems, which are parts of organizations, the more general findings above have been narrowed to the field of business information systems and analysed whether the same conclusions are valid.

Information systems are established in an organization in order to support the problem solving processes of it. It is obvious this way that the objectives of the system are in line with the goals of the organization. Formalized processes and workflow systems—which are also forms of information systems—aim to support the activities of the organization by defining business processes and automating them. Information systems operate in a problem space that originates from the organization and includes the same elements: organizational peculiarities, contextual and situational characteristics.

These attributes would suggest that the interpretation of effectiveness is also the same in information systems as in the case of organizations, meaning the degree to which objectives of the system are achieved.

However, it has to be noted that although the objectives of the information system is aligned with the goals of the organization, they are not the same. In the literature of management information systems, there were several attempts to measure information system effectiveness. A major part of them deduce system effectiveness from user satisfaction and system use, indicating that the measurement of the "goodness" of an information system is the level the users are satisfied. Although these approaches have also been criticised, major part of the discussions were about the selection of appropriate metrics for measuring user satisfaction. This is however a very complex problem where a lot of sophisticated approaches have been introduced from the field of cognitive and behavioral sciences for example.

Although there is no consensus on the adequate measuring of user satisfaction, factors of it can be obviously reasoned. Most important implication is that the user—who is working in an organization to solve problems assigned to him or her—is more satisfied when his or her work burden is reduced or

he or she is capable of solving more problems in a given time, or in other words, efficiency.

The implication of this finding is that efficiency can contribute to the effectiveness of the system. It also has to be noted however that improving efficiency can only be considered as a way of augmenting effectiveness when the scope of the system is not narrowed, or in other words, the system reaches the same goals as before. In terms of information overload, when the overload situation is alleviated at a given point in a business process, the burden on the user is decreased, improving his or her efficiency this way. When decreasing information overload is done a way that does not change the objectives of the system in a negative direction, it can be stated that the effectiveness of the system is improved.

III.1.2 Workflow

The operation of a system (including organization in this interpretation) is of highly dynamic by nature. In order to support the operation of the organization on a basic level, it is important to clarify the difference between business processes and workflow. Processes are realizing one function in an organization, consisting of activities following each other, possibly including decision points, forks and joins. Activities usually have inputs and outputs. During the execution of a process, some kind of transformation is performed on the inputs.

Workflow in contrast is a tool of process execution on one hand, or a series of processes on the other hand. Business processes are generally defined separately, are loaded into a workflow engine, which then controls process execution according to business needs. This way, workflow always executes process instances, moreover it transects them by executing multiple processes according to parameters or environmental settings.

Workflow also provides the necessary inputs for the activities in the processes, realizing this way a push approach to information delivery. In the current knowledge intensive world, it is rare that the provided information are sufficient enough to perform the task, it is likely that search for further information is involved, which is the pull approach of information delivery. On one hand, this pull method is served by appropriate background systems, while also supported by guidelines explaining how and where to look for the necessary information. This more-or-less ideal case is however only available in well-organized systems and organizations and can be considered as the "good" approach to information pull.

The "bad" approach is when the push support is minimal—e.g. when no workflow is present or the supplied information is not sufficient—therefore

the operation of the organization is heavily based on habitual routines developed by the employees that is not formalized and recorded at all in most of the cases. This situation however can be still valid when new procedures are emerging in an organization, for example in the case of changes in the business processes. It is imperative in this case that after the more-or-less proper processes are crystallized as best practices, they have to be recorded as formalized processes. This is one of the most powerful ways of conserving organizational knowledge.

When the processes are formalized, the role of push method increases while the importance of information pull diminishes. As changes in the environment of an organization are quite frequent, the ratio of information push and pull always changes, which makes their balance dynamic. The ever iterating transition between the pull and push approach can also be considered as a way of organizational learning.

III.1.3 Optimizing information push and pull

In order to find an appropriate balance between information push and pull approaches, a mathematical calculation is demonstrated based on the work of Kiss (1979). Workflow is transformed into a graph by noting decision points as vertices. All of the edges represent transitions between vertices, having a certain probabilities. In the beginning, these probabilities are equal, since the system is in a state of indetermination. Information can be added to the system either by executing the process—since being in a certain step shifts the probabilities of reaching other steps in the workflow—or adding information to decision points. The latter way can shift probabilities because uncertainty is decreased (or at least not increased, according to Shannon 1948). It can be seen this way that information available at the decision points can be influenced by information push or pull.

Workflow can be considered partially a human, partially a mechanized system. In this environment, information can be obtained in two ways: either by finding it by humans or delivering it by the machinery.

This statement has two consequences:

1. employing human and mechanical components always involve costs; and

2. acquiring information can be realized by employing pull and push approaches.

It can be considered therefore that the utility of information at a given decision point depends on the probability of possible transitions to further

steps (the matrix of indirect transition probabilities is used to represent transition probabilities between two arbitrary decision points Kiss 1979, p.87), and the cost of acquiring the information.

There are two ways to provide information at the decision points: letting the humans find it or delivering it to them. Employing pull approach means letting human workers find the necessary information. Finding information is a nondeterministic task that depends heavily on the previous experiences and knowledge of the workers that means that finding the necessary information cannot be assured in every case. When not enough information is obtained before a decision then the decision itself can be sub-optimal, that can lead to a kind of loss originating from the lack of information. Moreover information processing capabilities of human beings are very limited. This means that the more information can be processed the more processing cost is involved in the form of lost productivity, stress or other means that can be monetized as costs.

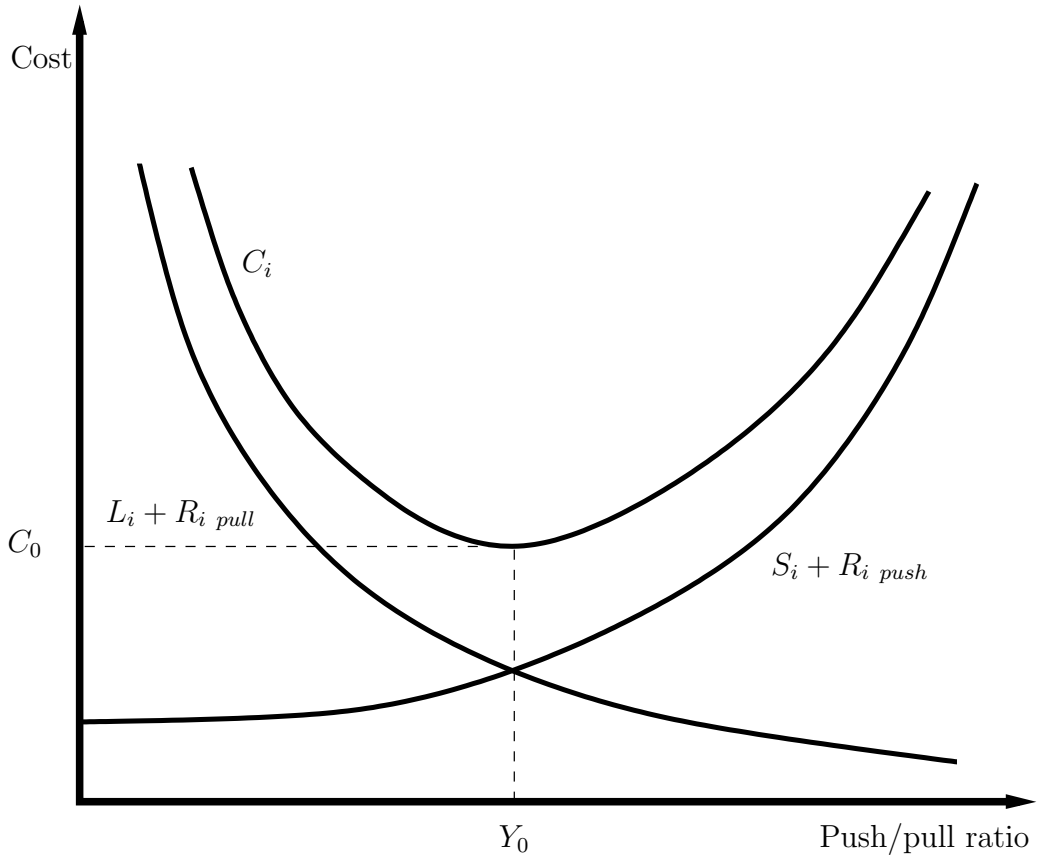

Figure 1: Costs of acquiring information

On the push side, delivering information requires computing, processing that can cause costs as computing power is involved on one hand and the sophisticated design and realization of data and information storage and retrieval facilities that are suitable to effectively handle and filter information. Considering ontologies as an example: the more sophisticated, detailed an ontology is, the more effort is involved in designing, implementing and maintaining it, which will be discussed later. Moreover—similarly to the pull case—pushing more information to the decision maker increases the cost of processing that information on the human side.

As a summary, it can be stated that costs can be defined this way:

$$c_i = \underbrace{L_i + R_{i\ pull}}_{\text{cost of pull}} + \underbrace{S_i + R_{i\ push}}_{\text{cost of push}} \tag{1}$$

where i is a certain decision point (vertex) in the graph, L_i is the loss from the lack of information—which is basically a kind of opportunity cost—, S_i is for the costs of push and finally $R_{i\ push}$ and $R_{i\ pull}$ represent the cost of processing information delivered by push and obtained by pull methods.

Regarding the analysis of costs, well-developed economic models can be applied. As push and pull approaches influence costs in the opposite direction, it is plausible to depict them on a figure (Figure 1) and it seems analogous to the economic theory of cost functions that is deeply discussed for example in Varian (1992, p.64–80).

In the case of our model, horizontal axis represents the push-pull ratio, which can basically be reduced to the amount of information push used. Push and pull complement each other in this system, meaning that the increase of push decreases pull and vice-versa. The curve of $L_i + R_{i\ pull}$ shows that by increasing information delivery for a certain decision by using push approach decreases the need of obtaining information by the pull method, which results in decreasing costs on the pull side. It can be easily seen that this is analogous to the short-run average fixed cost function depicted by Varian (1992, p.65).

Similarly, the function $S_i + R_{i\ push}$ increases when the amount of push is raising, since delivering more information requires greater processing capacity on the system side (represented by S_i) and on the human side as well ($R_{i\ push}$). It can also be seen that this is very similar to the short-run average variable cost function in the work cited above. The reason this function intersects the vertical axis above zero level is that workflow—which is the framework of our analysis—is by itself a form of information push, which involves certain costs compared to a disorganized situation. Function C_i aggregates the costs of information push and pull, resulting in a U-shaped curve matching the average cost function in microeconomic theory.

It is shown therefore that tools of microeconomic analysis can be applied to the cost of obtaining and delivering information, which implies that the same consequences can be drawn. According to the deduction of Varian (1992, p.65, 69), optimal choice of the amount of push is at the minimum of the C_i function that is denoted by Y_0, having a cost level of C_0. At this level, it is ensured that the result of both approaches take the minimal cost level. It is important to emphasize the assumption that information is equally useful, regardless its delivery method. This means that information acquired through manual searching (pull) and delivered by a system (push) is semantically equivalent, offers the same usefulness to its users. Information overload can only be interpreted when information processing does have certain limits or capacities, exceedint some ot these in terms of processing. Regarding a workflow system, capacities include factors like deadlines, costs, results and resources. An organization necessarily has a certain amount of these factors available. Considering the assumption detailed above, it can be concluded that the optimum ratio of information push and pull denoted by Y_0 offers the minimum of information overload in the organization.

III.2 Realizing information push in an ontology-based system that is appropriate for setting the ratio of push and pull

Other major achievement of the dissertation was the realization of an ontology-based system that can be tuned to set the appropriate ratio of push and pull by controlling the push side.

First, it is shown that structuring information or content is absolutely a necessity when they are to be sought in an information system. There are various tools and methods available for realizing this task, all of which having benefits and drawbacks.

As it was shown before, effectiveness highly relates to objectives of an organization. Regarding a system, operational environment includes the organization, meaning among others that the goal of the system is to support the objectives of the organization. Considering the balance of information push and pull, the same type of environmental information are also needed in order to be able to interpret the concepts and their relations. As it was seen, balance between push and pull approach is dynamic in time, ratio of push and pull changes according to the activities of the organization. There is, however a balancer that aims to strengthen the push side, organizational learning. The state of balance in a given time, as well as organizational learning can only be interpreted within certain boundaries that are provided

by the business context, the organization and the situation. These factors define a domain, and as Vas (2007, p.11, 14) has shown, ontologies can be built on specific domains

In summary, it can be stated that ontologies are appropriate tools for realizing information delivery by applying push method. Ontologies can be used to define a structure that not only contains concepts of a given domain, but their relationships, which can contribute to the preciseness of the queries and reasoning. Moreover, application of ontologies is scalable in terms of efforts, since the granularity of the concepts and relationships (e.g. logical statements) can be chosen according to the needs of the system (in terms of requirements of queries) or the organization. Naturally, developing a more fine-grained ontology requires more effort, which is realized in an increased cost level, resulting in a monotonously increasing cost function that fits our previous findings.

As realizing a system is not feasible in a general manner, a case study is employed in order to define the environment of the system. This pilot case is the determination of the framework numbers of Hungarian Higher Education, which is a knowledge-intensive process involving a lot of information processing steps. This pilot case was used in the research project called SAKE[1], in which the author has taken great part regarding the development of a semantically enhanced content management system and designing the semantic layer and information integration.

Solution to the pilot problem is offered at two levels. On the first level, systems are introduced that are able to support the process on a level that is currently available. These systems are a content management system responsible for storing pieces of contents (or documents); a groupware system for facilitating collaboration among the participants; and a workflow management system that governs the business process execution. On the second level, all of these information available in those isolated systems in various forms and amount, are integrated by using an ontology-based solution. According to the details of the case, ontology models have been developed in order to define the problem space of the pilot case. Information ontology contains details about the various information sources and forms of information stored in the system. Actor ontology describes actors and their roles, while Event ontology captures specificities of events happening in the system, adding temporal dynamics into the model. Preference ontology is about de-

[1]SAKE – Semantic-enabled Agile Knowledge-based e-Government (IST 027128) research project was pursued by an international consortium of partners, and co-financed by the 6th EU Framework Programme for Research and Technological Development. The SAKE project commenced on the 1st of March 2006 and lasted for 36 months. For more information, please refer to http://www.sake-project.org.

scribing preference rules of the users for querying and reasoning, Process ontology is used to represent runtime and persisted states of the business processes. These general type of information could be applied in any kind of information system. As one of the most important models, a Domain ontology has also been developed, which only contains information about the specificities of the pilot case, realizing this way a real customization of the model. During the process of determining framework numbers of Higher Education, a number of educational outputs and job offers are gathered and analysed. Most important part of the process is the comparison of the two set of information, which is a highly knowledge intensive task and can be complicated for public administrators as well since huge amount of information should be filtered and organized. Educational outputs and labor market demand are grasped through competencies, which are formalized in the domain ontology.

The realized system itself offers the possibility of uploading documents, annotate them by ontology concepts (e.g. competencies), facilitate discussions in the groupware system and offering a framework for the process by employing the workflow system. All event that happens in the system are captured in a semantic way by creating instances of ontological concepts and establishing appropriate relationships between them. Finally, the logic stored in the ontology is used for reasoning, resulting in the matching of job offers and educational outputs based on their logical restrictions according the offered or requested competencies. It is shown that the granularity of ontology can be chosen according to the needs and cost considerations of the organization. It is also shown that the result depends only on the granularity of the ontology, algorithmic implementations can remain unchanged. The ontology-based system is eligible this way to set the amount of information push used.

III.3 Benefits and importance of the research

Main benefit and achievement of the research is that it offers a viable approach of handling information overload within the frames of organizational workflow processes. Most important practical achievement is the ontology-based system that realizes information push in a controllable way. Granularity—and hence the costs of push and pull—can be influenced by changing the granularity of the ontology, by introducing more detailed or less detailed design.

As another important achievement, an ontological framework has also been presented that describes general and specific information about the organizational environment. This general model and the way the information

integration has been realized with the help of them can be used by any other complex information system in a similar setting. Domain ontology, on the other hand, can be freely designed in order to adjust it to the given problem. This offers the system the maximum flexibility.

Flexibility by itself cannot be considered good or bad. The level of detailedness however should be determined according to the given problem. The third important benefit of this research is to offer a methodology to determine the appropriate level of granularity by introducing and optimizing the costs of information delivery.

III.4 Future research considerations

The interpretation of information overload was restricted in this work to the organization, more precisely the organizational workflow. Other relations of the concept was mentioned and reviewed in the theoretical foundations of the dissertation, which provide a lot of area where useful investigation could be realized. Other disciplines, like cognitive sciences could also be involved into the analysis. Releasing the relatively strict environment of organizational workflows can also provide additional benefits to further researches.

In the current research, information contained by content management systems, groupware and workflow systems have been involved, analysed and integrated. The ontology-based information integration discussed in the dissertation is able to offer a general-purpose solution, which could be developed into a general semantic integration solution that could integrate any kinds of systems that have appropriate adaptors to the semantic layer. Although this approach seems to be feasible, limitations have to be considered as well—in the case of achievements of this research as well. Developing ontologies is quite complex task that requires professional experience and a lot of efforts as well. Ontology engines (like KAON2 in the current case) require many resources for their operation in terms of computing time and operative storage capacities. Use of newer or more efficient reasoning engines or other kind of logic description formalism can push the development of information integration solutions further.

Improvements could be performed regarding the cost-based model of determining the optimal ratio of information push and pull. Best practices of calculating or estimating costs of information processing, potential loss caused by the lack of information or costs of ontology maintenance could be collected and synthesized. Other factors regarding the model could also be determined and incorporated.

IV Major references

Ackoff, R. L. (1989), 'From data to wisdom', *Journal of Applied Systems Analysis* **16**, 3–9.

Antal-Mokos, Z., Balaton, K., Drótos, G. & Tari, E. (2000), *Stratégia és szervezet*, KJK Kerszöv, Budapest.

Babbie, E. (1996), *A társadalomtudományi kutatás gyakorlata*, Balassi Kiadó.

Blaxter, L., Hughes, C. & Tight, M. (2006), *How to Research*, 3 edn, Open University Press.

BusinessDictionary.com (2009), 'effectiveness'.
URL: *http://www.businessdictionary.com/definition/effectiveness.html Accessed on 04.06.2009*

Corcho, O., Fernández-López, M. & Gómez-Pérez, A. (2003), 'Methodologies, tools and languages for building ontologies. where is their meeting point?', *Data & Knowledge Engineering* **46**, 41–64.

Demeter, K. (2006), *A kompetencia – Kihívások és értelmezések*, Országos Közoktatási Intézet.

Edmunds, A. & Morris, A. (2000), 'The problem of information overload in business organisations: a review of the literature', *International Journal of Information Management* **20**(1), 17–28.

Eisenhardt, K. M. (1989), 'Building theories from case study research', *Academy of Management Review* **14**(4), 532–550.

Eppler, M. J. & Mengis, J. (2008), The concept of information overload - a review of literature from organization science, accounting, marketing, MIS, and related disciplines, *in* M. Meckel & B. F. Schmid, eds, 'Kommunikationsmanagement im Wandel', Gabler, pp. 271–305.

Füstös, L., Meszéna, G. & Simonné, M. N. (1986), *A sokváltozós adatelemzés statisztikai módszerei*, Akadémiai Kiadó, Budapest.

Hart, C. (1998), *Doing a Literature Review – Releasing the Social Science Research Imagination*, Sage Publications, London.

Himma, K. (2007), 'The concept of information overload: A preliminary step in understanding the nature of a harmful information-related condition', *Ethics and Information Technology* **9**(4), 259–272.

Kiss, I. (1979), *Bevezetés a számítástechnikába II. (Az informatika alapjai)*, Tankönyvkiadó, Budapest.

Knerr, T. (2007), A framework for representing rule-based user preferences in the semantic web, Master's thesis, Hochschule Furtwangen.

Knerr, T., Ntioudis, S., Papadakis, A. & Stojanovic, N. (2006), 'D6 – the SAKE ontology representing the conceptual model of the knowledge-based e-government', SAKE Project documentation.

Oakley, A. (1999), Peoples way of knowing: gender and methodology, *in* S. Hood, B. Mayall & S. Oliver, eds, 'Critical issues in social research', Open University Press, pp. 154–170.

Papadopoulou, S., Papadakis, A., Knerr, T., Stojanovic, N., Bruhács, T., Kő, A., Gábor, A., Kovács, B., Siemek, A. & Butka, P. (2006), 'D5 – as is analysis', SAKE Project documentation.

Shannon, C. (1948), 'A mathematical theory of communication', *Bell System Technical Journal* **27**, 379–423, 623–656.
URL: *http://cm.bell-labs.com/cm/ms/what/shannonday/shannon1948.pdf*

Shenk, D. (1997), *Data Smog: Surviving the Information Glut*, HarperEdge, San Francisco, CA.

Stake, R. (1995), *The art of case research*, Sage Publications, Newbury Park, CA.

Stojanovic, N., Apostolou, D., Dioudis, S., Biró, M., Gábor, A., Kovács, B., Butka, P., Lukac, G., Siemek, A., Wozniak, D., Petrvalsky, E., Bruhács, T. & Kasprzycki, J. (2008), 'D24 – Integration plan', SAKE Project documentation.

Tellis, W. (1997), 'Introduction to case study', *The Qualitative Report* **3**(2).
URL: *http://www.nova.edu/ssss/QR/QR3-2/tellis1.html*
Accessed on 30.04.2009

Varian, H. R. (1992), *Microeconomic analysis*, 3rd edn, W. W. Norton & Company, New York.

Vas, R. F. (2007), Tudásfelmérést támogató oktatási ontológia szerepe és alkalmazási lehetőségei, PhD thesis, Budapesti Corvinus Egyetem, Gazdálkodástani Doktori Iskola.

Yin, R. K. (1994), *Case Study Research*, 2 edn, Sage Publications, London.

V Publications in the topic

V.1 Book chapters

Futó, I., Gábor, A., Kovács, B. & Kő, A. (2008), Higher Education Portfolio Alignment with World of Labour Needs, *in* Ferro, E., Scholl, J. & Wimmer, M., eds., 'Electronic Government - Proceedings of 7th International Conference EGOV 2008, Turin 1-4 Sept', Trauner Verlag, Linz, 265–272.

Kismihók, G., Kovács, B. & Vas, R. (2009), Integrating Ontology-based Content Management into a Mobilized Learning Environment, *in* Caballé, S., Xhafa, F., Daradoumis, T. & Juan, A. A., eds., 'Architectures for Distributed and Complex M-Learning Systems: Applying Intelligent Technologies', IGI Global, Hershey, USA, chapter 10 (under publication)

Kovács, B. (2008), Szemantikus technológiák a tartalomkezelésben, *in* Bartók, I. & Simon, J., eds., '60 éves a Közgazdaságtudományi Egyetem Tanulmánykötet', Aula Kiadó, 407–423.

Kovács, B. & Vas, R. (2008), Ontology-based Content Management Systems in Public Administration, *in* Schweighofer, E., ed., 'Legal Informatics - The LEFIS Series, Vol 2.', Prensas Universitarias de Zaragoza, 45–62.

V.2 Articles

Kovács, B. (2008), 'Improving content management - a semantic approach', *Acta Cybernetica* **18**(4), 579–593.

V.3 Studies

Papadopoulou, S., Papadakis, A., Knerr, T., Stojanovic, N., Bruhács, T., Kő, A., Gábor, A., Kovács, B., Siemek, A. & Butka, P. (2006), 'D5 – as is analysis', SAKE Project documentation.

Stojanovic, N., Apostolou, D., Dioudis, S., Biró, M., Gábor, A., Kovács, B., Butka, P., Lukac, G., Siemek, A., Wozniak, D., Petrvalsky, E., Bruhács, T. & Kasprzycki, J. (2008), 'D24 – Integration plan', SAKE Project documentation.