

G.J. Houben and F. Dignum. Information for organized work. In F. Baader, M. Jeusfeld, W. Nutt (eds), *Proceedings 4th Int. workshop Knowledge Representation meets databases*, Athens, 1997.

# Integrating information for organized work

G.J. Houben and F. Dignum

Faculty of Mathematics & Computer Science

Eindhoven University of Technology

PO.Box 513, 5600 MB Eindhoven, The Netherlands

tel. +31 40 2472733, fax. +31 40 2463992, email: {houben,dignum}@win.tue.nl

## Abstract

In this position statement we claim that the organization of information in databases, knowledge bases, hypertexts and other integrated structures depends on the processes for which they are used. These processes can be highly structured, completely ad hoc or anything in between. Whether a standard procedure can effectively and efficiently be followed in the process, however, depends on the quality of the information available. The human actors in the process need a mechanism to relate the quality of the information and the quality of the standard procedure to the goals of the business activity. We argue that agent technology can be used to bridge the gap between highly structured situations with high quality data and ad hoc situations where little information is available or the information is not of the right quality.

## 1 Introduction

One might argue that ten years ago the most important electronically available information sources in companies were the (relational) databases. At that time all information relevant to the processes of the company would be modeled in these databases. Besides these large databases there also existed some small knowledge based systems that provided information for a small amount of specialized processes that needed more complex information.

This situation, however, has changed drastically over the past few years. Many more forms of storage, dissemination and processing of electronic information have become available and are widely used in practice. This trend towards integrated information systems is illustrated by the attention of concepts like hypertext systems, Intranets, document management systems (e.g. Documentum), groupware products (e.g. Notes, Exchange), and workflow management systems (e.g. Staffware, Cosa).

While the software tools may be using traditional database technology, the users perceive that the tools enable them to use new ways of communication. For the users the data and information is stored and accessed in a completely different way as before.

The traditional perspective of data and information processing does not cover the essential aspects of modern integrated information systems. It is not sufficient anymore to look only at a uniform language to represent information in the traditional database and knowledge base formats, but one should look at ways to access many different types of information intelligently. The key aspects are the integration of information from different sources, and the differences in disclosure principles for information accessed through different tools.

We propose to tackle this problem starting from the organizational point of view instead of starting from a technological point of view. That is, the starting point of describing the

information should be the humans that need the information to perform some activity within an organization.

## **2 Quality of information**

The human users of the information systems figure as the key actors in the processes of an organization. When they perform their activities, they process data and information retrieved from information systems. Traditionally, the approach has been that the architecture of the information system reflects the data processing aspects of some business process: the designer of an information system has designed the system with a prescribed usage in mind. Therefore, the user is actually guided by the information system itself.

Nowadays much more unstructured information can be stored through hypertext, document information systems, etc. These systems are used as general knowledge repositories and are not focused on one process. This flexibility also has its consequences on the actual information processing. In a large number of organizations knowledge workers need mechanisms to deduce which type of information can be used to perform a business activity. More than before, the knowledge worker must decide on the data to be used, the sources where the data can be retrieved, the format in which the data is retrieved, and the general quality of the information produced. Because information is stored in many different ways, an assessment of its quality is necessary to determine its value for a business activity.

The quality of information is determined by a number of properties. Some of these properties deal with the functionality: “how (with which semantics) is the information used?”. Other important quality aspects include time (validity, availability), costs, and resources (human, information). In the end the quality is related to the effectiveness and the efficiency for the activity at hand.

Essentially the users lack knowledge:

- Knowledge about the quality of information available within the organization: in order to be able to use a truly integrated information system, users must possess knowledge (meta-information) about the business information; without knowing what the information is worth, they cannot produce an optimal product or service.
- Knowledge of the activities and processes in which they use the information: in the traditional dedicated information systems the users/system interaction reflects the intended processes; in most of the modern practical cases the processes have a goal of their own, not reflected in the specific information systems: in the abstract integrated information systems it is left to the knowledge worker to use the specific information systems in order to fulfill the process goals.

## **3 Process information**

An approach that in the last years tries to support knowledge workers in their grip on the processes that play a role in the integrated information systems, is the use of process models. Concepts like workflow management, document management or product data management, start from a view on the information that focuses more on the process, and less on the data.

Considering workflow management, we see that the information processing in an organization is perceived as a logistical process. In this process the data or information is simply considered as material that must be distributed among the right users. Moreover, the humans involved are seen only as resources. This paradigm is generally valid if the work processes are highly structured, and produce end-products in high volumes.

While the design of an optimal logistical process contributes heavily to the performance of the information processing, this approach alone does not quite acknowledge the problem of using information from different sources. The information used in the workflow originates in

different databases and knowledge bases, but also comes from the less formal knowledge of the workers involved.

This knowledge of the workers also includes information on the applicability of a predefined process (procedure) to a specific case. In ad hoc or less structured situations knowledge workers need to customize the procedure to the actual needs of the specific case. This type of knowledge and insight is not acknowledged in workflow applications.

## 4 Integration of process and data aspects

The integration of all relevant aspects can be tackled bottom up from the databases and knowledge bases. In the first two sections we have argued that there are a number of difficulties in the bottom up use of information. Especially, it is rather difficult to assess the quality of the information used, with respect to the activities on hand: “what does this information contribute to the product or service that I am producing?”.

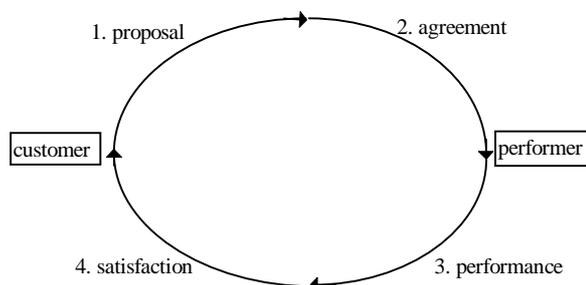
In the previous section we have argued that the top down approach suggested by the workflow management paradigm, is not always the route to follow. It does not acknowledge that the use of information requires knowledge (meta-information) on the effectiveness and efficiency of using that information.

We feel that a combined approach is necessary. Specifically in cases where the processes are less structured, may be even ad hoc, and where the quality of the information is an issue, we argue that an adequate support of the knowledge worker is necessary. They should have the knowledge to assess the combination of procedure and information that is available and relate that to the goal of the activities to be performed. By this we mean that the human actors have the means to decide how an activity is actually performed, while using their professional knowledge of the available data and of the prescribed procedure as a suggestion for an ideal process.

We aim at using agent technology to assist the users in navigating the different information sources rather than aiming at building a uniform interface to all sources.

## 5 Agent technology

We suggest an agent-based approach as basis to support the work of persons in an organization. It takes the persons as central entities in the workflow and models the workflow through commitments between the persons involved. The approach is called Action Workflow [ MWFF]. Each task can be modeled through the following cycle:

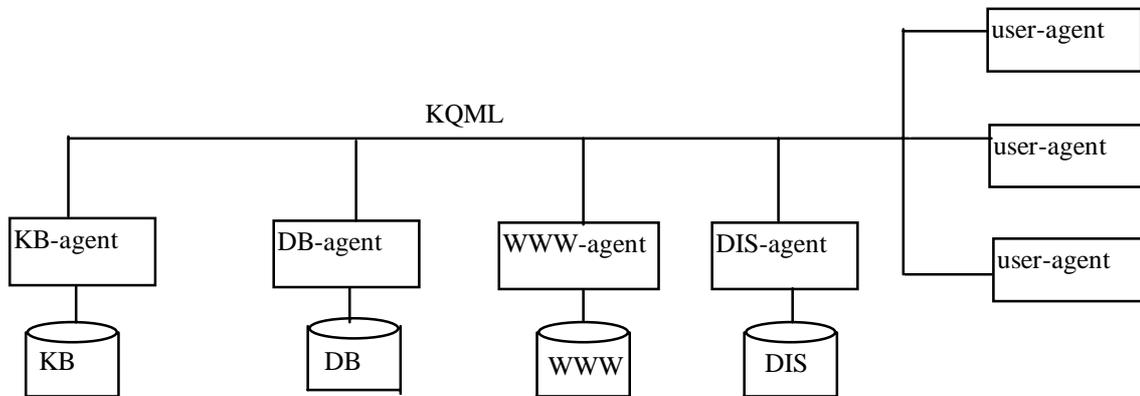


The third arc in this cycle indicates the step where the actual work is performed. We support this step by using a cooperative information agent as described in [VD]. In principle the agent performs two tasks. First it assesses whether the standard procedure is appropriate for the present task based on the goal, a formal description of the standard procedure and the quality of the available information. The second task of the agent is to collect the necessary information from the heterogeneous information sources. The information that is needed depends on the procedure that is followed: the standard procedure or an ad hoc solution that is

created for the task at hand. In order to perform its tasks the agent should have the following elements in its knowledge base:

1. Formal description of the standard procedure,
2. Type and location of necessary information (including persons),
3. Formal description of alternative steps to be taken,
4. Goal formation rules to plan the ad hoc procedure

For each information resource type there is an agent that can assist the user to access that type of information. The agent that supports the task of the user is called the user-agent, the other agents are called resource-agents. Their architecture is identical, but the user-agent contains knowledge about the task it supports and the location and types of information needed for that task, while the resource agents contain knowledge about the type of resource they maintain. They can translate the agent communication language into the language for that particular resource. As agent communication language we will use KQML, but we will use only a few performatives that we have predefined and for which a clear semantics is given. These performatives are things like, ASSERT, COMMIT, REQUEST, etc.



In practice the approach can imply that the agents help the humans to decide on the procedure to be used. While in some cases the agents can be developed to automatically assess the properties of the procedure and information, in most cases the humans will act as the users of the knowledge delivered by the agents. This means that the humans are in control and they combine their own professional expertise with the knowledge from the agents to decide on the procedure to follow in order to achieve the goal at hand.

## References

[MWFF] Raúl Medina-Mora, Terry Winograd, Rodrigo Flores, Fernando Flores,  
The Action Workflow Approach to Workflow Management Technology,  
in: Computer-Supported Cooperative Work 92 Proceedings, 1992, pg. 281-288.

[VD] Egon Verharen and Frank Dignum,  
Cooperative Information Agents and Communication,  
in: Cooperative Information Agents 1st Int. workshop, CIA'97, (LNAI 1202), Springer-Verlag, 1997, pg.195-209.