Data, Information, and Knowledge Modeling in Worksystem Networks

Sangay Dorji and Marite Kirikova

Riga Technical University, Faculty of Computer Science and Information Technology, Institute of Applied Computer Systems, Kalku 1, Riga, Latvia
sdorji753@gmail.com; marite.kirikova@cs.rtu.lv

Abstract. Data, information and knowledge are widely used concepts and often perceived as synonyms, but in reality there are situations when the differences between these concepts have to be taken into consideration. This paper examines a possibility to distinguish between data, information and knowledge flows in worksystem networks. An enterprise architecture frame that consists of several basic elements of ArchiMate language is proposed for distinguishing between data, information, and knowledge flows in worksystem networks.

Keywords: Worksystems, Data, Information and Knowledge Flow Analysis, ArchiMate.

1 Introduction

Distinguishing between concepts “data”, “information”, and “knowledge” mainly is the topic of academic discussions, however, also in the practice there are situations, when distinguishing between these three concepts is important. For instance, availability of data not necessarily means that the information needed for a particular employee really is acquired and used. To have the information, the employee has to know about the availability of the data and should be able to interpret the data, thus, s/he has to be informed about the data and has to have knowledge suitable for the interpretation of this data.

We have already analyzed data, information, and knowledge (DIK) flows separately in the context of Viable Systems Model [1], where the frame of particular elements of enterprise architecture modeling language ArchiMate [2] was proposed for DIK flow representation. However, that frame was applicable only to information flows in a single worksystem [3], as it did not include elements for the data transfer via physical networks. In the networks of worksystems, e.g., networks of enterprises or in enterprises that heavily depend on their internal communication via networks, the data transfer via physical networks is an important issue as it concerns such aspects as security, network availability, etc. Therefore, we have reexamined the frame proposed in [1] and compared it to other enterprise architecture usage approaches especially focusing on worksystem networks.

The paper is organized as follows. The basic concepts, namely, DIK flows, worksystem networks and ArchiMate language are briefly discussed in Section 2. Section 3 describes the method used for detecting the enterprise frame for DIK flows in worksystem networks. Section 4 discusses experiments for analyzing DIK flows in worksystem networks. In Section 5 the limitations of the study, conclusions, and future works are briefly outlined.

2 Basic Concepts

In this section the main concepts used in the paper are introduced on the basis of the related works. Data, information, and knowledge definitions are overviewed in Subsection 2.1. Worksystems are briefly discussed in Subsection 2.2. The reasons of the use of the ArchiMate language for the representation of DIK flows are explained in Section 2.3.

2.1 Data, Information, and Knowledge Definitions

There are lots of definitions of data, information, and knowledge presented by many researchers. Not always the authors define all three terms. For analysis we have selected the list of definitions where all three concepts are defined by one and the same author(s) [4,5,6,7] with the purpose to focus on the differences of the terms and to avoid similar definitions of different concepts in cases when authors do not distinguish between the concepts of data, information, and knowledge.

While there were differences in the statements on what data, information, and knowledge is in the selected definitions, still there were the following commonalities.

- With respect to data:
  - Data is collective pieces of values to produce distinct information and facts representing ideas, objects, or conditions.
  - Data is the basis of reasoning and calculations.
  - In order to create suitable information for any decision making, collective arrangements of data have to be properly analyzed.
  - Data is observed, un-interpreted symbols.
- With respect to information:
  - Information is details or facts learned about something or somebody through sequential arrangement of things.
  - Information is interpreted symbols and symbol structures.
  - Information is obtained after the analysis of properly collected data.
  - Information may lead to an increase in understanding and decrease in uncertainties.
  - Information is a key to any decision making, behavior, or an outcome.
- With respect to knowledge:
  - Knowledge is an understanding about a subject that is acquired through the experience.
  - Knowledge is gained through studying the range of accumulated information regarding particular subject.

– Knowledge needs extensive amount of experience focusing on the information related to the subject.

To distinguish between data, information, and knowledge we also used their descriptions from [8,9] and [10,11], and [12]. Here information is regarded as data interpreted by knowledge. From the data definitions we came to know that data is raw facts or un-interpreted symbols like words, numbers, characters, and signals, which do not provide any meaning. Once the data is analyzed and arranged properly, and it is possible to understand data and provide the meaning, then data will become information. To convert the raw fact or un-interpreted data to interpreted data we need the process. This process is called *data interpretation*. While interpreting the data humans have to make a decision based upon their past experience, observation, culture, and educational background to provide contextual meaning to data. Humans interpret data using their knowledge that performs the process of data interpretation. When it comes to computers, they need different algorithms to interpret the data.

Once the data is interpreted, it is transformed into information. This information has to be made understandable. Once it is understood by the users, they can justify what are the main causes and consequences, and what are the additional features, problems, or suggestion required. Basically, the elaboration is the kind of problem solving in worksystems. For instance, in the worksystem we can elaborate information like how many foreign students are registered in the university; whether all registered students have successfully completed their course or not; to attract foreign students what main courses are needed, etc.

The worksystems interact with the environment and tend to grow so that there will be lots of data, information, and knowledge in the organizations. Using elaborated information and interpreted data, worksystems can “learn” (based on past and present information) and predict the future growth of the worksystems. Learning is the integration process of new information from the existing information in worksystems and making decision for future. Thus, knowledge should be able to interpret data, elaborate information, and learn from interpreted data and elaborated information.

2.2 Worksystem Networks

A worksystem can be any kind of organization, which involves human participants to operate the machine or machine performing automatically on the basis of human instructions or built in algorithms, by using data, information and technology resources so that it will benefit both organization and the customers.

According to Steven Alter [3] *“A worksystem can be defined as a system in which human participants or machines perform work using information, technology, and other resources to produce products and services for internal or external customers”*. The author stated that the *“Customers and products/services may be partially inside and partially outside because customers often participate in the processes and activities within the worksystem and because products/services take shape within the worksystem. Processes and activities, participants, information, and technologies are viewed as completely within the worksystem. Environment, infrastructure, and strategies are viewed as largely outside the worksystem even though they have direct* *and indirect effects within the worksystem”* (see the graphical representation in Figure 1).

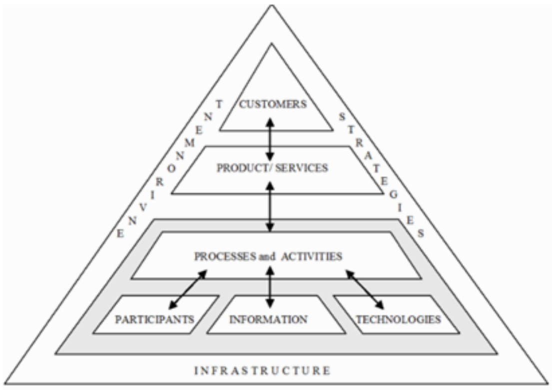

Fig. 1. Framework of worksystems (adopted from [3])

According to Steven Alter [3], the worksystem framework consists of the following nine elements (see also Figure 1):

- *Customers*. Customers are those people who benefit or receive products and services directly from a worksystem. They are the ones who can use or experience the quality of the products and services of the worksystem. The customers can be either external or internal customers.
- *Products and Services*. Products and services can be physical products, information products, and services made by a worksystem for various customers. Examples of products and services are arrangements, agreements, goods, consultations, etc.
- *Processes and Activities*. Processes and activities involve detailed steps of work in the worksystem. The sequence or details of doing work in a worksystem depend on individual skills, knowledge, experience, and observations which help to make decisions, communicate with clients, and coordinate the work in the worksystem.
- *Participants.* Participants are those people who directly or indirectly are involved in performing the work in a worksystem. For instance, an employee directly is involved in the worksystem to perform the work; but the customer, who demands the product and services to the worksystem, is not performing the work, but still is considered as an external participant because of contributing something to the worksystem.
- *Information*. Information can be categorized into two parts that is codified and un-codified information. The information can be manipulated (created, updated, deleted) by using processes and activities of the worksystem. We can see that the worksystem’s framework here does not distinguish between data, information, and knowledge.
- *Technologies*. Technologies are tools involved in a worksystem that help the employees to perform the work easily. There are two types of technology, the one which is operated by employees and another one that performs the work autonomously.

- *Infrastructure.* An infrastructure consists of human, informational, and technical resources that a worksystem relies upon, but which are outside the worksystem and are shared resources with other worksystems. Human infrastructure is the people and organizational units that supply services shared by different worksystems, for instance, training organization. Information infrastructure is information shared across various worksystems; it can be shared database and other enterprise information. Technical infrastructure includes hardware and software which helps worksystem to share the information between multiple worksystems. An example of technical infrastructure is an Enterprise Resources Planning (ERP) suite.
- *Strategies.* Strategies are some kind of worksystem's guidelines which help the organizations to achieve their primary goals. Strategies can be worksystem strategy, departmental strategy, and enterprise strategy.
- *Environment.* An environment is viewed largely outside a worksystem and it needs to follow respective laws, standards, culture, policies, and regulations.

Worksystem networks can be defined as the collection of human employees, computers, servers, network devices, and other peripherals that are connected with each other to communicate, exchange the information, and share hardware and software resources for mutual benefits. Internet is the network of networks where the information can be exchanged globally. There are other options, such as local area networks and metropolitan area networks which can help the worksystem to exchange data, information, and knowledge within the worksystem network. Inside a worksystem, worksystem networks are important because they help the worksystem and its sub-worksystems to share data, information, and knowledge between the nodes based on certain rules and principles. In a worksystem, knowledge holders can be actors, roles and application components, such as, e.g., websites. Data can be text, symbols, images, etc.; interpreted data can be regarded as information. (See Figure 2).

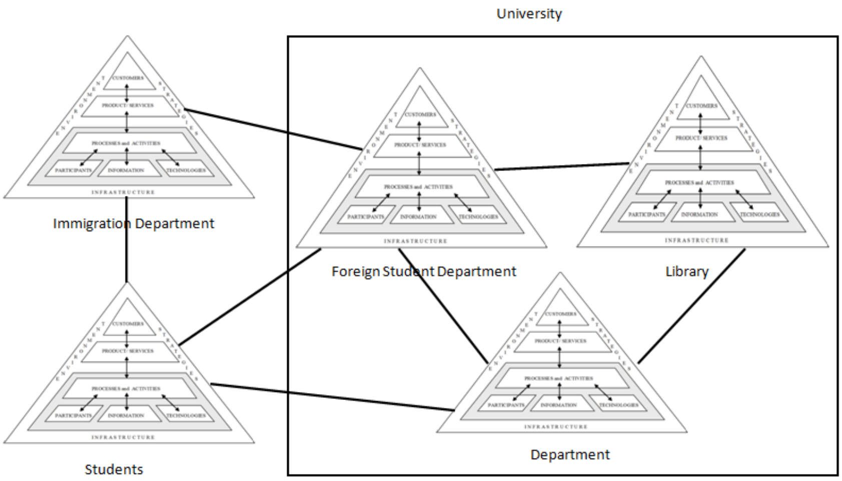

Fig. 2. Example of the worksystem network (University is represented as a worksystem consisting of the network of worksystems)

2.3 ArchiMate

ArchiMate initially, in 2002-2004, was developed in The Netherlands by a project team from Telematica Institute in cooperation with Dutch partners from government, industry, and academics. It is an Open Group Standard, modeling language for enterprise architecture. It is a visual language to represent end-to-end enterprise architecture in terms of business processes, applications, and technology [2], [13,14,15]. The core of the language consists of three layers. The Business layer mostly describes the business processes and people (called business actors) involved in the business processes. Business actors can be humans, departments, and business units. They may be individuals or groups of people. Each actor is assigned with a business role. This level shows, how the business events, processes, services, and functions are related among themselves and to the associated individual business units. Information, product, process, and organization domains should be included in this layer. The Application layer consists of application components, i.e., application software that performs the particular tasks. In other words, information is processed by application software. It supports the business layer with application services, which are realized by application components. The Technology layer mostly deals with infrastructure services like processing, storage, and communication infrastructure, needed to support specified applications in the applications layer. The Technology layer refers to the technical infrastructure domain.

The ArchiMate language consists of three types of elements. There are *Active elements* that represent those elements in the real world that "exhibit behaviors". *Behavior* displays the actual behavior that can be observed in the real world. The examples are business processes, application services, and infrastructure services. *Passive structure/ passive elements* are also called as information. They represent those things that undergo or are the result of the behavior. These are elements that cannot act and which are acted upon by that behavior [13].

The following are some of the advantages of using ArchiMate language for representation of DIK flows in worksystems:

- It has elements for all concepts discussed in Subsections 2.1 and 2.2.
- It is an independent enterprise architecture modeling language.
- It is Open group standard and is supported by free, cross-platform tools to create ArchiMate models.
- It is easy to understand by experts and non-experts across all domains.
- It is able to visualize the relations between the domains; and gives a possibility to visualize the models in different ways, e.g., we can view the model like a business process, a product, an application usage, an application structure, and an infrastructure.

3 The Method

We have already analyzed DIK flows in a single worksystem [1], [11,12] and created the enterprise architecture frame for this purpose. To find the frame for worksystem networks we proceeded as follows.

1. We analyzed related works with the purpose to find alternative frames for representation of DIK flows (one alternative was selected).
2. On the basis of the DIK flow representation frame developed in our earlier work [1], [16,17], and the selected alternative frame, we draw a hypothesis with respect to the representation of DIK flows in worksystem networks.
3. We created worksystem network scenarios and represented each data, information, and knowledge flow in them separately.
4. On the basis of acquired representations we selected the frame for DIK flow analysis.

In related works we found a university meta-model from Coventry University [18] that reflected all elements relevant for worksystem networks. This model was made using the ArchiMate language. We modified the model so that it can represent two worksystems simultaneously (we repeated elements from the center to the right also from the center to the left, i.e., added these elements to the original meta-model (see Figure 3).

By extending the Coventry University model we achieved its similarity with the enterprise architecture frame used for DIK flow analysis in a single worksystem [1] reflected in Figure 4.

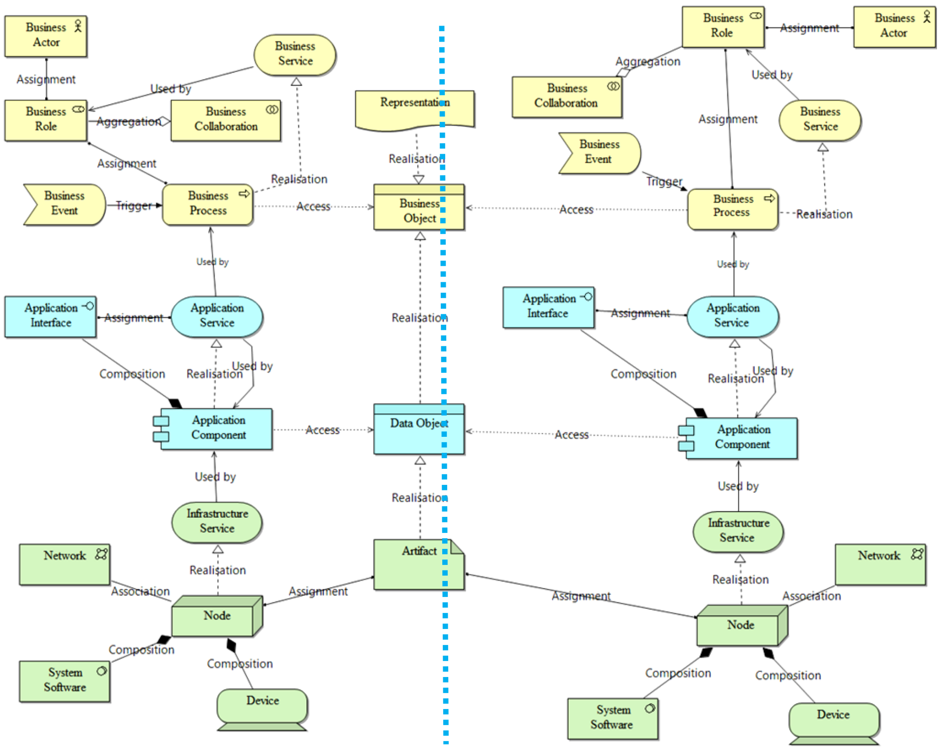

Fig. 3. Extended University Meta model (original available in [18])

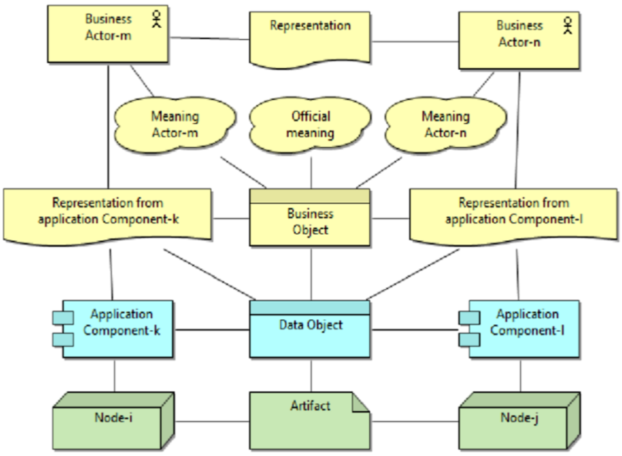

Fig. 4. Generic enterprise architecture construct used for flow analysis (adopted from [1])

When comparing the models (frames) in Figure 3 and Figure 4, one can see the similarity in their central part. On the basis of this similarity the hypothesis was drawn that for representation of DIK flow this central part of the frame is essential, and that, for the flow to exist, an element of the central part has to be connected with at least one element on the left of the center and at least one element on the right of the center. The hypothesis is partly illustrated in Figure 5.

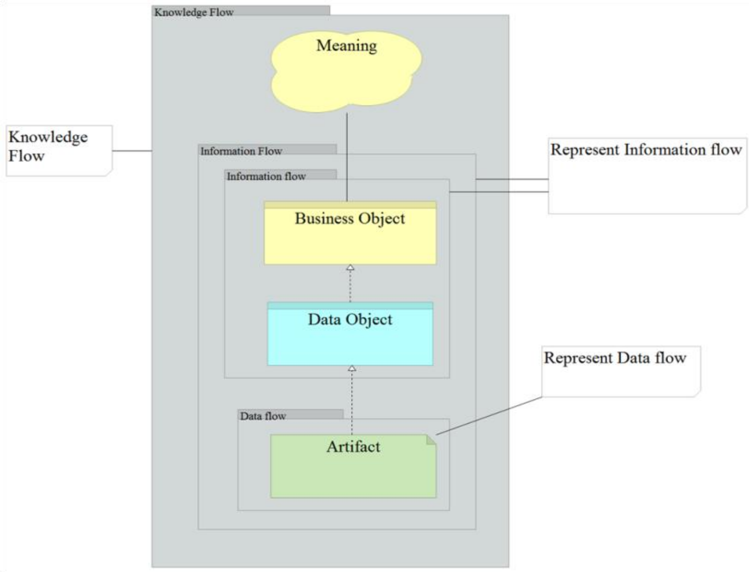

Fig. 5. Hypothesis of DIK flow

We assumed that in ArchiMate we can represent the one-step DIK flows where in the bottom level we have an artifact, which represents the file, a piece of data, or message which can be shared and transferred to other nodes; so it represents the data flow. In the second layer we have a business object and a data object.

The data object is realized on a business object, for example the 'bank account' Business Object may be data (a Data Object) of an accounting application; it is the same item's representation in a different architectural layer. Data object is the other way of representing the business object in the application layer. It represents the information flow.

There are two ways to represent the information flow, - one way is using the data object realized on a business object, and another way is when artifact is realized on a data object and then realized to a business object. It will become knowledge flow, if there is a meaning attached to the business object - which should enable the receiver to interpret data, elaborate the information, and also learn from the information.

4 Experiments with Scenarios

To experiment with DIK flow representations, we have created the following five scenarios based on the Riga Technical University Foreign Students Department activities:

- Student searching and applying for getting admission in the university;
- Borrowing books from the university library;
- Participating in sport membership at the university;
- Student taking academic leave;
- Student wanting to retake the examination for the second chance.

For each scenario all DIK flows were mapped into the enterprise architecture frames to see which elements are involved in the representation of the flows. Altogether representations of 32 DIK flows were created. In the representations all the block elements in the right column and the left column of the frames were considered as two worksystems, and hypothesis elements (Figure 5) were represent between them. This helped to distinguish data, information, and knowledge flow between the worksystems.

The DIK flows between two worksystems were shown in the respective diagrams with numbers which represent the flow propagation. For instance, in Figure 6 the excerpt of the scenario on getting admission in the university is represented. Once the University IT department gets the information from the student, the IT department employee can provide a temporary user name or password for the access to the Virtual Learning Environment (VLE) Ortus by the email. According to the hypothesis, as the flow contains the artifact, data object, and business object, it is an information flow. So the elements of the frame that are related to the numbered lines are necessary to represent the information flow. In the same manner other data, information, or knowledge flows found in the scenarios were represented.

After representing data, information, and knowledge flows and comparing the representations, it was concluded that it is sufficient to have the elements of the generic frame illustrated in Figure 4, but it is very important to add a network component (available in Coventry University meta-model, see Figure 3). Generic flow analysis frame represented in Figure 4 does not show exactly the flow analysis in physical worksystem networks. If we add a network component then it is possible to show how the data, information, and knowledge are exchanged between two or more worksystem networks. This is illustrated in Figure 7 where the reconstructed flow representation frame is presented.

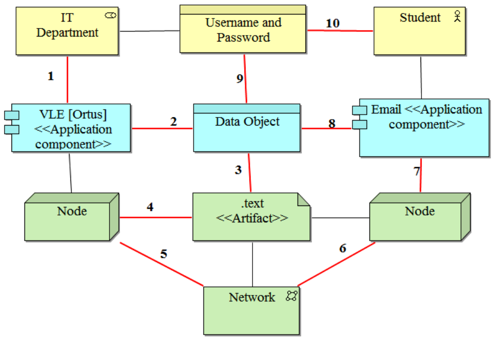

Fig. 6. IT departments provides the username and password to the student

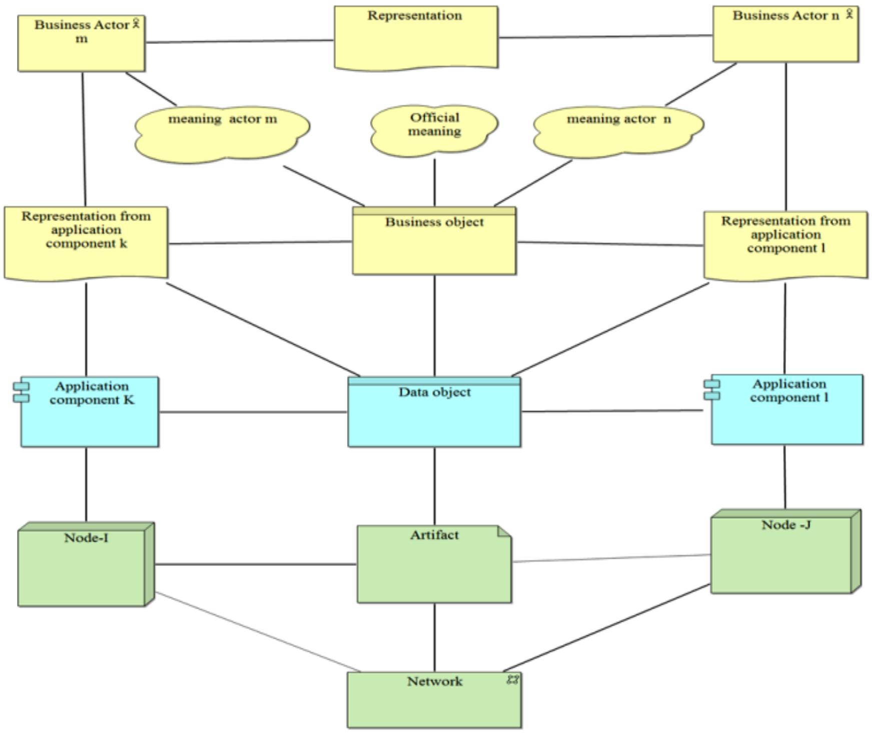

Fig. 7. Reconstracted enterprise architecture frame for DIK flow analysis

While in the most of cases the reconstructed frame fulfilled the hypothesis of how to distinguish between data, information, and knowledge flows, there were also such situations where it was complicated to use the construct presented in Figure 7. These were situations where worksystems of different granularity were exchanging the information, i.e., the worksystem of the one side was of the smallest possible granularity (e.g., student) but another worksystem consisted of several other worksystems (e.g., a department, software applications, computers). Nevertheless, the frame was sufficient to cover all 32 DIK flows and did not require additional elements, such as business event or service that were represented in the alternative model (Figure 3).

The experiments with the scenarios showed that distinguishing between the data, information and knowledge flows in worksystem networks requires specific methods and usually their scrupulous consideration is not a part of information systems development activities. In this experiment we represented only one step flows. For multistep flows there should be additional methods and algorithms developed that can check the consistency of the flows. However, even the one step flow analysis gives an opportunity to examine DIK flows closer and, for instance, to check whether really the information is received not just data is available or to see whether there is knowledge to be transferred to enhance the data interpretation or learning.

5 Conclusion

Today the information systems really influence the worksystem or organization to do the business and move forward to achieve their goals by helping the management to carry out daily operations and control and monitor their progress. To benefit from the information systems it is important to distinguish between data, information, and knowledge to ensure that decision makers are equipped with needed knowledge information, and data. Also, it has to be mentioned that data, information, and knowledge may require different treatment from the point of view of storage, reusability, security, and other issues of information systems management.

This paper illustrates a step forward to distinguishing between these three phenomena in worksystem networks. While the ArchiMate language gives an opportunity to represent all elements of worksystem framework, only part of the language elements were needed for the scenarios that represented data, information, and knowledge flows in everyday activities of the foreign students department of the university.

This study has several limitations: (1) only the one-step flows were represented in the proposed frame; (2) the differences between data, information, and knowledge flows were not discussed in detail; (3) in the representations it was not always possible to handle information fusion present in the scenarios; (4) there were difficulties to represent flows between worksystems of different granularities; and (5) only 5 scenarios and 32 flows were analyzed.

Future research is intended for multistep DIK flows and broader range of scenarios.

Acknowledgment. The research reflected in this paper was supported in part by the Latvian Council of Science, grant for project No 342/2012.

References

1. Kirikova, M., Pudane, M.: Viable system model based information flow. In: New Trends in Databases and Information Systems. Advances in Intelligent Systems and Computing, Volume 241, Part 1, pp. 97–104. Springer (2014)
2. Welcome to ArchiMate 2.1, an Open Group Standard, The Open Group, 2012–2013. pubs.opengroup.org/architecture/archimate2-doc
3. Alter, S.: Worksystem Theory: Overview of Core Concepts, Extensions and Challenges for the future. In: Journal of the Association for Information Systems, vol. 14, no. 2, pp. 72–121 (2013)
4. Zins, C.: Data, Information, Knowledge. In: Knowledge Map of Information Science (2013)
5. Bellinger, G., Castro, D., Mills, A.: Data, Information, Knowledge, and Wisdom. In: System Thinking, A journey in the realm of systems (2004)
6. Frost, A.: Defining Knowledge, Information, Data (2013)
7. Zins, C.: Conceptual approaches for defining data, information and knowledge. In: Journal of the Association for Information Science and Technology, vol. 58, no. 4, pp. 479–493 (2007)
8. Aamodta, A., Nygardb, M.: Different roles and mutual dependencies of data, information and knowledge – an AI perspective on their integration. In: Data and Knowledge Engineering, vol. 16, pp. 191–222 (1995)
9. Genkin, M.: What's the difference between data, information and knowledge in machine systems? Or is there a difference? https://www.quora.com/What%E2%80%99s-the-difference-between-data-information-and-knowledge-in-machine-systems-Or-is-there-a-difference
10. Zhung, H.: Knowledge flow network planning and simulation. ScienceDirect, pp. 571–592 (2005)
11. Yamin, M., Otto, J.: Patterns of knowledge flows and MNE innovation performance. In: Journal of International Management, pp. 239–258 (2004)
12. Rajalakshmi, S., Banu, R.W.: Analysis of Tacit Knowledge Sharing and Codification in Higher Education. In: International Conference on Computer Communication and Informatics, Coimbatore (2012)
13. Gils, B.V., Dijk, S.V.: Experiences, techniques, and best practices. In: The Practice of Enterprise Architecture, BiZZdesign Academy (2013)
14. Sobolewska, B., A.: When the applications are business process. The Netherlands, University of Leiden (2011)
15. Kock, N.: System analysts and design fundamentals: A Business Process Redesign Approach, SAGE Publications, Inc. (2006)
16. Kirikova, M.: Work Systems Paradigm and Frames for Fractal Architecture of Information Systems. CAiSE Forum 2014, Thessaloniki, Greece, June 16-20, 2014, Selected Extended Papers. Information Systems Engineering in Complex Environments, Vol. 204, Lecture Notes in Business Information Processing, pp. 165-180. Springer (2015
17. Kirikova, M.: Modeling for Viability. Perspectives in Business Informatics Research. Vol. 229 of the series LNBIP, 14th International Conference, BIR 2015, Tartu, Estonia, August 26-28, pp. 265-272 (2015)
18. Czechowski, N., Padam, S., Anderson, I., Woodcock, C.: JISC Enterprise Architecture and ArchiMate Project., Coventry University, London (2011)