PROVIDING SEMANTIC INTEROPERABILITY FOR INTEGRATED HEALTHCARE USING A MODEL TRANSFORMATION APPROACH

Barbara Franz[(a)], Herwig Mayr[(b)]

[(a)] Upper Austria University of Applied Sciences, Research Center Hagenberg, Austria
[(b)] Upper Austria University of Applied Sciences,
Department of Medical Informatics and Bioinformatics, Hagenberg, Austria

[(a)]Barbara.Franz@fh-hagenberg.at, [(b)]Herwig.Mayr@fh-hagenberg.at

ABSTRACT
Integrated care and the achievement of high quality healthcare over institutional borders require technical and semantic interoperability between different healthcare providers. Thus, a meta model was developed, based on the application of a system conformant to Integrating the Healthcare Enterprise (IHE) and the use of Clinical Document Architecture (CDA) as document format. This meta model combines the properties of several health information systems and different health service domains (HSD).

For each HSD, a model can be derived, which describes used IHE profiles, available document types to the point of coding systems used in exchanged documents. Thus, it provides information about the domain, workflows, patient etc., which are used for the transformation into another domain. Using this model transformation, HSD models may be compared, checked and completed, where applicable. The results can then in turn be applied to the CDA documents. Thus, it helps to improve the communication among healthcare institutions.

The evaluation of developed models and transformations is conducted using genuine healthcare data, provided by e-Care, an IHE-conformant system for exchange of healthcare data between several healthcare providers.

Keywords: *model transformation, semantic interoperability, Clinical Document Architecture, Integrating the Healthcare Enterprise*

1. THE CHALLENGE OF HEALTHCARE DATA EXCHANGE
Efficient data exchange between healthcare providers is the cornerstone of integrated healthcare (Arrow et al. 2009). To achieve a high quality of healthcare over institutional borders, it is important to regard which activities and examinations have already taken place and in which quality they were conducted. Therefore it is necessary for healthcare providers to document in such a way that access to documented information as well as a comparison and an integration of the data into other systems is also possible for other providers.

This is a big challenge as the project e-Care, where several healthcare providers were connected, has shown (Franz et al. 2009a). The most demanding factors are:

- various electronic health information systems,
- different domain languages and understanding,
- special processes and workflows in various organizations,
- varying documentation with different goals,
- varying structures in documentation,
- different designations and
- little to no use of coding systems.

Thus, the challenge is the successful combination of technologies and information models in integrated healthcare. Part of this can be met by relying on Integrating the Healthcare Enterprise (IHE). The worldwide initiative IHE aims to optimize the interoperability of IT-Systems in medicine and healthcare, using international standards, cf., e.g., (HL7 2010a). IHE provides several so-called integration profiles which define use cases and suggest technical solutions (IHE 2010a).

One IHE profile suggests the use of Clinical Document Architecture (CDA) as a uniform format for the exchange of clinical documents (IHE 2010c). Therefore, current national and international projects like e-Care, ELGA (ELGA 2010), VHitG (VHitG 2007) and epSOS (epSOS 2010) boost the use of fully structured or at least semi-structured healthcare data in form of HL7 V3 CDA-documents. But as long as there are no strict requirements for system providers to provide their data in highly granular structures, loads of unstructured data will always exist which cannot be interpreted unambiguously (Wozak et al. 2008).

Since uniform semantic standards are only used rudimentarily in these documents (as described above) it is seldom possible to directly compare, check (on plausibility, correctness and completeness) or integrate document content. Such an integration is only possible in consideration of semantics and domain specific differences. Using semantic and technical interoperability in various health service domains, documented information has to be homogenized in a way that an access to structured as well as unstructured information, consistent with a certain context, is possible. Such a knowledge transfer over institutions demands a clear and structured language (Perhab 2007) and particularly interoperability.

2. TECHNICAL AND SEMANTIC INTEROPERABILITY

National as well as international initiatives for information integration in healthcare aim at the increase of interoperability of information systems and at minimizing integration efforts (Norgall 2003; Sunyaev et al. 2008). The term *interoperability* denotes the ability of systems for collaboration. *Technical interoperability* indicates the interaction of technical components and systems (Heitmann und Gobrecht 2009). To achieve interoperability of healthcare systems, IHE and also Continua Health Alliance (CHA 2010a) provide efficient methods and standards.

Semantic interoperability, on the contrary, denotes the interpretation of data, while preserving the intended meaning. It can be achieved using common information models and uniform terminology, thus enabling cross-community interpretation, processing and storage. Besides technical standards, an agreement is essential, how medical and domain specific terminology is used. Nevertheless, maintenance as well as further development and optimization of the terminologies and their structural elements should not be underestimated (Heitmann und Gobrecht 2009).

Domain specific terminology has to be used unambiguously to prohibit misunderstandings and to improve collaboration. Only this way semantic interoperability can be achieved (at least partially) in one domain as well as across several domains.

2.1. IHE Integration Profiles

The use of IHE, conformance to IHE integration profiles and adherence to suggested standards are one way towards smooth integration (IHE 2010b). Four of the most important integration profiles, which enable interoperability, are:

- *Patient Identifier Cross-Referencing (PIX)* supports the unique identification of patients with different patient identities in several domains. For this purpose, patient data are sent from a source to a Patient Identifier Cross-Reference Manager (IHE 2010a).
- *Patient Demographic Query (PDQ)* defines a central patient register, which allows distributed applications to query demographic and case-related patient information (IHE 2010a).
- *Cross Enterprise Document Sharing (XDS)* supports the patient care of several healthcare providers by allowing registration, distribution and access of documents referring to one patient. XDS is content neutral, i.e. it supports different types of documents, independent of their content or format (IHE 2010a), although other profiles force the use of certain standards for documents.
- *Cross Enterprise Document Workflow (XDW)* focuses on the management of cross-enterprise healthcare workflows using a specific workflow document that references all documents related to a clinical workflow and manages changes in document states (Zalunardo and Cocchiglia 2011).
- *Patient Care Coordination (PCC)* describes the cross-enterprise sharing of medical summaries and personal health records (IHE 2010c).

2.2. Health Level Seven

IT standards, as suggested by IHE integration profiles, are required for data exchange between different information systems. In healthcare, the most important standards, besides Digital Imaging and Communications in Medicine (DICOM) for the exchange of images, are Health Level Seven (HL7) with Version 2 and 3 as well as CDA, which allow the exchange of information like documents and messages.

HL7 standards have been specifically developed for the health sector. They define the exchange of messages, document based communications as well as cooperating services, their implementation and necessary infrastructural services (HL7 2010a). Two important HL7 types for this work are:

- *HL7 Version 2*, which is primarily used in hospitals for message exchange between well-established systems,
- *HL7 Version 3*, which is based on the Reference Information Model (RIM, c.f. Fig.1) and on the Extended Markup Language (XML) and is used for trans-sectorial message exchange in the entire health sector.

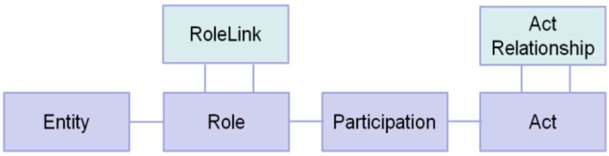

Figure 1: Core classes of the RIM

The formal computational exchange format for HL7 artefacts is defined by the *Model Interchange Format* (MIF), which is a part of the HL7 Version 3 methodology documentation (McKenzie et al. 2011). To represent the information content required to support a particular domain within HL7 and thus provide interoperability for a specific domain, a *Domain Information Model* (DIM) can be derived from the RIM.

The RIM is a generic healthcare specific information model. Figure 1 shows a part of the HL7 RIM, i.e. the four core classes and the two most important additional classes of the model:

- *Act*: e.g. an observation or treatment,
- *Entity*: e.g. a person,
- *Role*: e.g. author,
- *Participation*: relationship between act and role, e.g. a doctor performs a treatment,
- *ActRelationship*: relationship between acts, e.g. a diagnosis results from an observation,

- *RoleLink*: relationship between roles, e.g. organizational hierarchy.

The goal of HL7 Version 3 is the development of a uniform understanding of objects and processes in the healthcare environment. The use of RIM provides specifications to structure, type, content as well as semantics, used vocabulary and underlying processes necessary for data transfer and interoperability.

Besides the exchange of messages, HL7 Version 3 also provides CDA as specification for the structure, content and exchange of clinical documentation. CDA is based on the RIM and defines the structure and content of medical documents using XML. A CDA document consists of a structured header and a structured or unstructured body. The header contains information about the document, patient, patient encounter, participants and relations to other documents, guidelines or templates. The body contains the actual content of the document.

As already mentioned in section 1, CDA is exerted in several developments and projects, e.g. the European project epSOS, where it is the specified document format. These projects and developments specify CDA guidelines, which describe the underlying models and therefore the structure and parts of the semantic of the documents and their content. Connections between classes and attributes in the model are shown and possible restrictions on the CDA schema (XML schema) specifications are detailed. Partial structures of the CDA which are defined as CDA templates can be identified over a template ID.

CDA templates are predefined models which simulate the structure of documents or parts thereof and of data elements. Other guidelines are partially based on existing templates, which are referenced. Using the template OID in CDA documents, an automatic technical conformity-check against guidelines is possible. Examples for templates which are often referenced are Continuity of Care Document (CCD) and parts of the IHE profile PCC.

2.3. Use of Code Systems

The unambiguous use of domain specific terminology is a means towards semantic interoperability. In the health sector, different code systems are used, which provide a uniform definition of terminology (Heitmann und Gobrecht 2009). Examples for such code systems are:

- International Statistical Classification of Diseases and Related Health Problems (ICD),
- Logical Observation Identifier Names and Codes (LOINC),
- Systematized Nomenclature of Medicine – Clinical Terms (SNOMED-CT),
- International Classification of Functioning, Disability and Health (ICF),
- Medical Dictionary for Regulatory Activities (MedDRA),
- Nationality and language codes (ISO tables),
- Unified Code for Units of Measure (UCUM).

Since such code systems are not always used and fully structured information can therefore rarely be provided, unstructured and only partly structured (semi-structured) information have to be taken into account. Therefore, the various types of information have to be connected to provide semantic interoperability, which is necessary for comparing and checking structured as well as semi-structured data.

Current research like (Kilic und Dogac 2009; Bointner und Duftschmid 2008; Rinner und Duftschmid 2009) only consider finely structured CDA data, whereas other work in this research area like (Spat et al. 2007; Faulstich et al. 2008) are limited to fully unstructured free text documents. Also, the comparison of models in one domain has already been successful, as described in (Franz et al. 2009b). However, preserving semantic interoperability is required not only in one domain, but across several domains.

3. A MODEL BASED APPROACH USING IHE AND CDA DATA

Our approach presented in this paper can be used for structured data as well as free text and semi-structured information in form of CDA documents which are exchanged across several domains.

Based on the application of an IHE conformant system and the use of CDA as document format a meta model has been developed which combines the properties of several health information systems and different health service domains (HSD). The HSD meta model contains information about IHE profiles, for example which meta data are available about documents and patients etc. HSD models are derived from this meta model.

A HSD model describes which metadata is actually available in the domain, e.g. which IHE profiles are used, which document types are available, how CDA documents are structured to the point of which coding systems are used in the exchanged documents. Thus, we have information about the domains, workflows, patient etc., which are used for the transformation into another domain.

Derived domain models build the foundation for a transformation between different models and for the cross-domain integration of healthcare data. They can be compared, checked and completed, where applicable. The results can then in turn be applied to the CDA documents.

3.1. Analysis

In the analysis process the systems and healthcare service environments which are considered were identified and it was defined how to use the relations between these systems and domains. Furthermore, information for detailing the domain models like coding systems and similarities in the structure of CDA documents is recorded, to set the semantics into context using a domain specific description.

To achieve this, an analysis of domain knowledge, for example about the processes and terminology of a specific healthcare domain, was necessary and forms the basis for adequate methods and techniques for the model transformation. During a hospital stay, for example, various (CDA) documents are produced (diagnoses, medical summaries, care summaries, operation reports,…). Analyzed IHE profiles provide metadata about these documents as well as detailed patient data, metadata about organizations, domains and workflow information.

Health service domains, inductive connections and especially national and international CDA guidelines were reviewed and checked for similarities, for example Patient Summary (epSOS), Austrian nursing care summary (ELGA) and the Continuity of Care Document (CCD). In all analyzed guidelines, CDA bodies are to be structured at least into sections. What stands out is that the CDA body sections as well as their content in CDA body sections are coded differently (for example LOINC-codes vs. national codes), even if containing equal information. Besides the use of HL7 vocabulary in the CDA header for certain data like gender, marital status and religious affiliation, suggested coding systems vary in general.

3.2. Modeling

The modeling process is based on the analysis results: Domain specific properties of the health service environments are represented in form of a meta model (Kühne 2006), which is – in this case – an abstraction of domain specific models. A domain is in this context a health service environment, which comes with a particular semantic, i.e. with particular terminology.

Each domain provides specific information which is taken into account when deriving a specific health service environment model for a domain. Since the HL7 DIM covers only structured data and requires a detailed definition of processes and data in each domain, building or deriving an explicit DIM from the RIM is not sufficient for practical use in this case. Experience shows that use of terminology and coding systems, which would provide detailed and unambiguous definitions, is only rudimentarily implemented in current practice.

Therefore, more information has to be taken into account and an extended meta model is built in which specifications are embedded like terminology-, coding systems, guidelines, templates, standards and IHE profiles. This results in the use of not only detailed but also rough definitions as domain specific constraints for deriving health service environment models (c.f. Fig.2), e.g.

- which IHE profiles are applied,
- which standards are implemented,
- which guidelines and templates systems, documents etc. are in conformance with, and
- which terminology and coding systems are used.

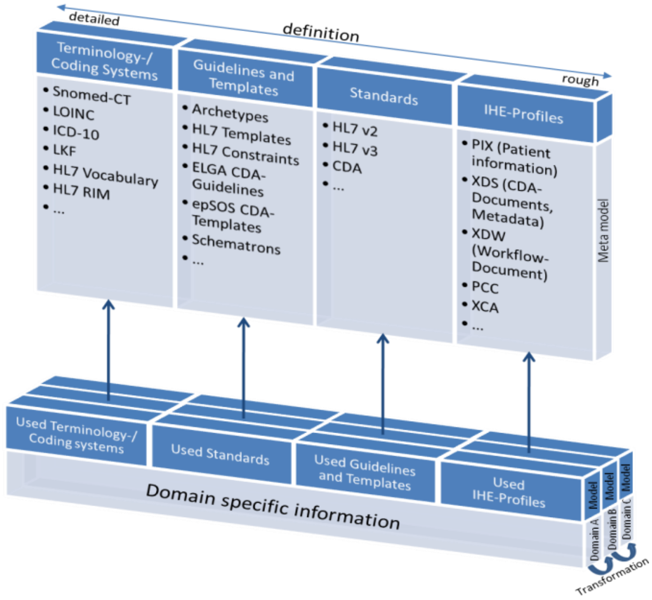

Figure 2: Modeling approach

Using the structure of CDA documents based on the RIM and document metadata provided by a system conformant to IHE recommendations, health service domain models for structured, non-structured and semi-structured data are derived from a meta model.

4. RESULTS AND EVALUATION

The derived HSD models predict content of documents and their relations, based on the particular domain. A specific document of a certain domain describes observed properties. These properties, for example, describe the occurrence of specific terms in a text or in a certain section in a CDA document, which can be deduced by a causal knowledge base.

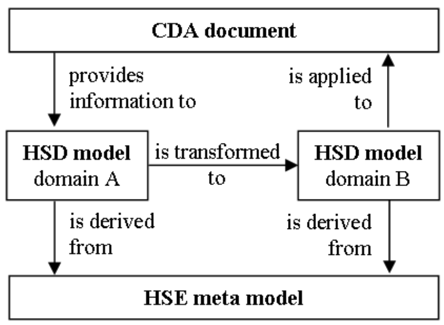

Figure 3: Model transformation

These domain models may be compared, checked and completed, where applicable, by the use of model transformation. The results can then in turn be applied to the CDA documents (c.f. Fig. 3). Thus, it helps to increase the availability of these documents and to improve the communication among healthcare institutions.

Due to known disadvantages of classic knowledge based systems (compare (Giarratano und Riley 2004)), the support of such a system using data driven technologies (Xu und Hou 2009) is focused. Additional technologies to describe the models and their properties, for example MIF, which is more complete than the XML representation (McKenzie et al. 2011), and an ontology-based approach combined with archetypes as suggested for example by (Duftschmid, Wrba and Rinner 2010), are tested.

The evaluation of the developed models and transformations is currently conducted using our e-Care system. The e-Care system (cf. Fig. 4) is an IHE-conformant platform for the exchange of healthcare data between several healthcare providers (Franz et. al 2009a). It was developed from 2008 to 2010 and has been established since 2010 in an entire region in Upper Austria for the exchange of documents between two clinics, five mobile nursing services and six nursing homes. Therefore, we are able to analyze the practicability of this research in a real-time environment using genuine healthcare data.

Figure 4: e-Care: exchange between hospitals, nursing homes and mobile nursing services

Especially the amount of semantic preservation in an IHE-conformant exchange of healthcare documents is analyzed and how feasible contents can be compared and checked for plausibility.

As plausibility checks cannot be done automatically, the collaboration with users, i.e. healthcare providers, is necessary. We plan to develop further methods and tools to enable health professionals to describe the static and dynamic structure of their health service environment and allow the mapping of one predefined domain into another by themselves.

5. CONCLUSION AND OUTLOOK

This research supports integrated care by enabling comparison, evaluation and completion of documents of various healthcare service environments. Thus, it helps to increase the availability of these documents and to improve the communication among healthcare institutions.

The model transformation approach is not only based on HL7 CDA but also uses metadata provided by an IHE infrastructure, which provides more domain information. Therefore, semantic interoperability may be preserved for highly structured data as well as semi-structured information across domains.

Since the model transformation may not always be automatically achieved, the knowledge of domain experts is necessary. Hence, a specific user interface has to be provided which enables healthcare providers on the one hand to manage the model transformation when necessary and on the other hand to work efficiently with model transformation results.

Over time, a large amount of documents is accumulated in an IHE-conformant system, especially for multimorbid patients. A next step might be to adapt the models so that they improve and refine themselves in a self-learning way using new documents for a patient. Thus, the interaction with domain experts during the transformation process could be reduced.

ACKNOWLEDGMENTS

This project is supported by the program Regionale Wettbewerbsfähigkeit OÖ 2010-2013, which is financed by the European Regional Development Fund and the Government of Upper Austria.

REFERENCES

Arrow, K. et al., 2009. Toward a 21st-Century Health Care System: Recommendations for Health Care Reform. *Annals of Internal Medicine*, vol. 150 no. 7 493-495.

Bointner, K., Duftschmid, G., 2008. Semantische Interoperabilität im elektronischen Gesundheitsdatenaustausch mittels dualer Modellierung – der HL7 Templates Ansatz. In Schreier, G., Hayn, D., Ammenwerth, E., eds.: *Proceedings e-Health2008 & e-Health Benchmarking 2008*, 29-30 May 2008, Vol. 235, Wien, OCG Books.

Continua Health Alliance, 2010. *About the Alliance*. Available from: http://www.continuaalliance.org/about-the-alliance.html [accessed 1 September 2010].

Duftschmid, G., Wrba, T., Rinner, C., 2010. Extraction of Standardized Archetyped Data from Electronic Health Record Systems based on the Entity-Attribute-Value Model, In: *International Journal of Medical Informatics*; Vol. 79(8); pp. 585 – 597.

ELGA GmbH, 2010. ELGA. *Elektronische Gesundheitsakte*, Available from: http://www.arge-elga.at [accessed 11 August 2010].

epSOS, 2010. *European Patients Smart Open Services – epSOS*, Available from: http://www.epsos.eu [accessed 11 August 2010].

Faulstich, L. C., Müller, F., Sander, A., Kreuzthaler, M., Kaiser, S., Errath, M., 2008. Semantisches Retrieval medizinischer Freitexte, *Proceedings 53. Jahrestagung der Deutschen Gesellschaft für Medizinische Informatik, Biometrie und Epidemiologie*. Stuttgart, 15-19 September 2008, Düsseldorf, German Medical Science GMS Publishing House

Franz, B., Lehner, M., Mayr, H., Mayr, H., 2009a. e-Care – IHE-Compliant Patient-Data Exchange in Order to Enable Home & Mobile Care of Elderly People Within an e-Care Affinity Domain. *Proceedings 6th International Conference on Information Technology: New Generations*, 25-27 April 2009, Las Vegas.

Franz, B., Mayr, H., Mayr, M., Pfeifer, F., Altmann, J., Lehner, M., 2009b. Integrated Care Using a Model-Based Patient Record Data Exchange Platform. *Proceedings 21st European Modeling and Simulation Symposium*, 23-25 September 2009, Tenerife, Spain.

Giarratano, J. C., Riley, G., 2004. *Expert Systems: Principles and Programming*, 4. Ed., Boston, Thomson/PWS Publishing Company.

Heitmann, K. U., Gobrecht, K., 2009. *HL7 Kommunikationsstandards für das Gesundheitswesen. Ein Überblick*, HL7 Benutzergruppe in Deutschland, Köln, LUP AG Lithographie & Printproduktion.

Health Level Seven, 2010a. *HL7. Health Level Seven*, Available from: http://www.hl7.org [accessed 1 September 2010].

Health Level Seven, 2010b. *HL7 V3 Clinical Document Architecture, Release 2.0*, ANSI Standard CDA Rel. 2.

Health Level Seven, 2010c. *HL7 Reference Information Model*, Available from: http://www.hl7.org/Library/data-model/RIM/C30201/rim.htm [accessed 3 November 2010].

IHE International, 2010a. *IHE Information Technology Infrastructure – Technical Framework, Volume 1 Integration Profiles*, Available from: http://www.ihe.net/Technical_Framework/index.cfm#IT [accessed 10 September 2010].

IHE International, 2010b. *IHE Information Technology Infrastructure – Technical Framework, Volume 2 Transactions*, Available from: http://www.ihe.net/Technical_Framework/index.cfm#IT [accessed 10 September 2010].

IHE International, 2010c. *Patient Care Coordination. Technical Framework*, Available from: http://www.ihe.net [accessed 10 September 2010].

Kilic, O., Dogac, A., 2009. *Achieving Clinical Statement Interoperability Using R-MIM and Archetype-based Semantic Transformations*, IEEE Transactions on Information Technology in Biomedicine, Vol. 13, No. 4, S. 467–477.

Kühne, T., 2006. *Matters of (Meta-)Modeling. Journal on Software and Systems Modeling*, 5(4)369–385, Berlin–Heidelberg, Springer.

Norgall, T., 2003. Kommunikationsstandard für die Gesundheitstelematik: Status und Ausblick. *Tagungsband e-Health 2003: Telematik im Gesundheitswesen – Vernetzte Versorgung*, Dresden.

Perhab, F, 2007. Babylonische Sprachverwirrung? Taxonomien, Klassifikationen – gemeinsame Pflegesprache und Wissensmanagement. In: *Österreichische Pflegezeitschrift*. 3/2007, pp. 8–16., Available from: http://www.oegkv.at [accessed 11 August 2010].

Rinner, C., Duftschmid, G., 2009. *Validieren von auf Zweimodell-Ansätzen basierenden, vollstrukturierten EHR-Daten am Beispiel EN/ISO 13606 und HL7 CDA*. In Schreier, G., Hayn, D., Ammenwerth, E., eds.: *eHealth 2009 – Health Informatics meets e-Health: Tagungsband e-Health 2009 & e-Health Benchmarking*, Vol.. 250, Wien, OCG Books.

Spat, S., Cadonna, B., Rakovac, I., Gütl, C., Leitner, H., Stark, G., Beck, P., 2007. *Multi-label Text Classification of German Language Medical Documents*, Health Technol. Inform. 129:1460-1.

Sunyaev, A., Schweiger, A., Leimeister, J.M., Krcmar, H., 2008. Software-Agenten zur Integration von Informationssystemen im Gesundheitswesen. In Bichler, M., Hess, T., Krcmar, H., Lechner, U., Matthes, F., Picot, A., Speitkamp, B., Wolf, P., eds. *Multikonferenz Wirtschaftsinformatik*, 26-28 February 2008, Berlin, GITO-Verlag.

VHitG, 2007. Arztbrief auf Basis der HL7 Clinical Document Architecture Release 2 für das deutsche Gesundheitswesen Version 1.5, Berlin.

Wozak, F., Ammenwerth, E., Hörbst, A., Sögner, P., Mair, R., Schabetsberger T., 2008. IHE Based Interoperability – Benefits and Challenges. In Schreier, G., Hayn, D., Ammenwerth, E., eds.: *Proceedings e-Health2008 & e-Health Benchmarking 2008*, 29-30 May 2008, Vol. 235, Wien, OCG Books.

Xu, J., Hou, Z., 2009. Notes on Data-driven System Approaches. In Tan, T. S., eds. *Acta automatica sinica*, 35(6)668–675, Beijing, The Chinese Association of Automation and the Institute of Automation, Chinese Academy of Sciences.

Zalunardo, L., Cocchiglia, A., 2011. *IHE IT Infrastructure (ITI) Technical Framework Supplement, Cross-Enterprise Document Workflow (XDW).* Available from: ftp://ftp.ihe.net/IT_Infrastructure/iheitiyr9-2011-2012/Technical_Cmte/Profile_Work/XDW/IHE_XDW_v16_23_03_11_draft.doc [accessed 5 April 2011]

AUTHORS' BIOGRAPHY

Barbara Franz MSc, is scientific researcher at the Research Center Hagenberg of the Upper Austria University of Applied Sciences.

Dr. Herwig Mayr is professor for software engineering at the Faculty for Informatics, Communication and Media of the Upper Austria University of Applied Sciences in Hagenberg.