Finnish National Archive of Health Information (KanTa): General Concepts and Information Model

● Teemu Suna

The need for semantically interoperable health information systems is an eminent principle in health informatics. The Service Orientated Architecture (SOA) together with a standardised structured information model based on the extensible markup language (XML) like HL7 CDA R2 is a recognised tool for storing and sharing medical information in a semantically uniform way (HL7: health level 7, CDA: clinical document architecture, R2: release 2). In Finland, the Ministry of Social Affairs and Health initiated an implementation project to build a national, centralised health information archive based on HL7 v3 CDA R2 in 2007 (v3: version 3). The Finnish model using the SOA architecture and a highly structured implementation of CDA R2 provides a solid infrastructure for gathering nationwide health information in the centralised archive. In Finland, these architectural foundations are applied to implement a national health record archive, a national electronic prescription system, and a Web portal for citizen access to personal health information. Together, these three applications form the foundation of the Finnish national archive of health information known as KanTa. Fujitsu Services Oy in Finland is the prime contractor and has overall project responsibility for delivering the KanTa system.

1. Introduction

In Finland, the public primary healthcare services are provided through municipal health centres. People use their local health centre because it has been the intrinsic principle of the Finnish healthcare system since 1972. For specialised healthcare, Finland is divided into 20 hospital districts.[1] All municipal health centres must be part of a hospital district to guarantee specialised healthcare services at an accessible distance. Moreover, because of the good availability of primary healthcare, it is only on rare occasions that patients need to be transported to a health centre in another municipality. Thus, it can be said that the Finnish healthcare system is mainly geographically organised.[2]

However, particularly when it comes to specialised healthcare, cooperation between healthcare professionals, regardless of geographical location, becomes more important, as many of the methods and therapies applied require distinct medical competencies. As medical experts can be located all over Finland, the ability to share the same background information about patients to support treatment choices is vital. Therefore, for collaboration, one of the most important tools is seamless access to the complete health information of patients.[3]

In Finland, most patient records are in electronic format and stored in locally acquired electronic health record (EHR) systems. Because of the geographically organised healthcare, EHR systems typically contain comprehensive patient records. However, EHR systems are mainly based on different software products and implemented

FUJITSU Sci. Tech. J., Vol. 47, No. 1, pp. 49–57 (January 2011)

49

by various system integrators. Consequently, they are mostly semantically incompatible and provide poor integrability, so cooperation between healthcare service providers is rarely based on online EHR exchange.[3] Moreover, the need for collaboration between healthcare professionals is expected to increase rather than decline, while interoperable health information systems are widely seen as an eminent principle in health informatics.[4]–[7]

To produce easily accessible semantically interoperable patient health records and a national concept for archiving health information, the Ministry of Social Affairs and Health of Finland initiated an implementation project to build a national centralised health information archive based on HL7 v3 CDA R2 in 2007 (HL7: health level 7, v3: version 3, CDA: clinical document architecture, R2: release 2). As an active archive, the system provides reliable up-to-date health information to the healthcare professionals participating in patient care.[8]

The Finnish model for the national health information archive builds on the Service Orientated Architecture (SOA) and on a highly structured implementation of CDA R2 documents and deploys a semantically interoperable infrastructure to gather nationwide health information in the centralised archive.[8],[9] In Finland, these principles are applied to implement a national health record archive, a national electronic prescription (ePrescription) system, and a Web portal for citizen access to health information. Together, these three applications are the first applications of the Finnish national archive of health information known as KanTa.[8] ePrescriptions will be the first application in clinical use, and practical use of this application is expected to begin progressively in 2010–2011.[10]

The Ministry of Social Affairs and Health of Finland is coordinating and the Social Insurance Institution of Finland (KELA) is responsible for implementing the KanTa project. On the basis of a public tender in 2007, KELA selected Fujitsu Services Oy in Finland (hereinafter, Fujitsu Services Oy) to be the prime contractor for delivering the KanTa system. Fujitsu has overall project responsibility for designing and building the KanTa system based on the guidelines for the national architecture specified by the Ministry of Social Affairs and Health of Finland.

This paper first describes the general concept of the initial parts of KanTa in order to give an overview of the system. It then describes the general information model used in KanTa. Here, the main objective is to overview semantic interoperability and the role of the HL7 CDA R2 standard in KanTa. The paper also aims to promote further discussion of nationwide semantically interoperable health information archiving and related possibilities, for example, at the personal level to enable follow-ups for health and prospective disease risks as well as at the organisational level to enable effortless collaboration within all healthcare institutions.

2. General concept of KanTa

The term KanTa is used to collectively describe the national healthcare information systems in Finland. In addition to the first three applications of KanTa, there are other development and implementation projects, such as the pharmaceutical database that will also be part of the KanTa system. The KanTa architecture defines the centralised information model, software design principles, and technical baseline for the national healthcare information systems in Finland.[8]

In the KanTa concept, the most elementary principle is to store and manage national healthcare information in a centralised system. To access the centralised system, KanTa provides a unified messaging interface based on HL7 v3 for integrating the systems.[11] Using that interface, EHR systems, pharmacy systems, and citizens using the health information portal can read and manage health information stored in

KanTa, depending on their access rights.

In Finland, the development of solutions in sharing electronic healthcare information has a long history. Regional solutions have led the way in cross-organisational healthcare information exchange in Finnish healthcare since 1998. However, the regional solutions are based on a reference register model. In other words, a reference to a patient record in a local EHR system is stored in a centralised register. A centralised secure Web portal reads the reference register and provides links to patient records stored in the local EHR systems. The reference register contains sufficient metadata to search for patient records from the centralised portal.[3]

The reference register model has achieved positive results in supporting regional cooperation between healthcare professionals.[12] However, despite the fact that users can access health information in another municipality's EHR system, the reference register model preserves the healthcare information in the proprietary structure specified by each EHR system, so healthcare information stored in the distributed systems remains semantically incompatible.[3]

In KanTa, semantic interoperability is achieved through a centralised information model and storage. Healthcare information stored in KanTa must follow the rules of the HL7 CDA R2-based information model of KanTa. That is, when KanTa receives information from the EHR system, the document inside the received message is initially validated first and stored only if the data is valid. As a result, the meaning and context of information stored in KanTa is unified, so it is semantically interoperable, which enables the information to be exploited.

2.1 First applications of KanTa

Implementation of the nationwide healthcare information archive began with the three main building blocks of the KanTa system: the national EHR archive (eArchive), the national ePrescription system, and the Web portal for citizen access to personal healthcare information.[8]

The eArchive provides healthcare professionals with centralised access to nationwide electronic patient records and treatment data. The EHR archive plays a central role in supporting the collaboration between healthcare organisations during treatment. It also provides means for referring back to health information histories and prevents duplicate or unnecessary procedures. While the health records are stored in a uniform format (HL7 CDA R2), the information is accessible from different healthcare systems using KanTa's HL7 v3 messaging interface.

The ePrescription system plays a central role in the KanTa system. To generate an ePrescription, a doctor issues and signs a prescription using the EHR system, which stores it in the ePrescription system, where it is held for 30 months. Patients can give their doctor permission to review their medication regimen in the centralised archive to prevent the issuing of duplicated drugs and unwanted complications. The ePrescriptions are stored and accessed in a manner similar to that used for eArchive.

The Web portal for citizen access to personal healthcare information is a Web application that provides a view of a citizen's prescription information and patient records. Citizens can see only their own healthcare information, i.e., the view is restricted to show only personal information of the authenticated user. Users can authenticate by using an online banking code or electronic identity (ID) and print out a summary of their ePrescriptions and see their patient records if such a record has been approved by a medical professional. Through the Web portal, ePrescriptions are accessible for 30 months from the prescription's issue date. Information stored in eArchive is generally accessible for the remainder of the patient's life.

FUJITSU Sci. Tech. J., Vol. 47, No. 1 (January 2011)

51

2.2 Overview of KanTa architecture

The KanTa systems are based on SOA, which is a concept for organising and utilising capabilities that are distributed and possibly under the control of different owners.[13] In KanTa, SOA is used to exploit well-known software development principles such as modularity, reusability, and scalability to implement a loosely coupled architecture. In other words, a collection of software modules is organised into independent entities and encapsulated as services to provide the business logic requirements of the KanTa system. By orchestrating service calls to be executed in a certain order, one can achieve the desired functionalities. For example, in KanTa, logic for authentication, archiving, and logging is encapsulated into separate services and called only when needed in the health information archiving process. The services are based on Web Services technology, and selected services are exposed to integrated systems through KanTa's messaging interface. Each Web Service is defined using the Web Services Description Language (WSDL) and the related XML Schema Definition (XSD), as specified in the Web Services Interoperability Organization's Basic Profile 1.1 standard. Thus, WSDL describes the Web Services and how to access them from the technical viewpoint, and XSD specifies the legal building blocks of the messages including the CDA R2-based structure of the healthcare information that can be stored in KanTa.[14] These architectural principles allow KanTa to be a modern and highly interoperable platform for national healthcare information archiving.

In this paper, KanTa's systems, actions, and roles are described at the general level. That is, individual services and systems are grouped on the basis of their functionality and role. On the basis of this grouping, the services and systems are divided into providers and consumers from the perspective of the KanTa system, as illustrated in **Figure 1**[8), 11)] and explained below.

2.2.1 Service consumers

The systems that use the KanTa services can be divided into three main groups.
1) The EHR systems of hospitals and other healthcare providers use KanTa to store and search health records and prescriptions. Health records are produced in the EHR systems and transferred to KanTa for archiving. When a healthcare professional needs health information from the centralised archive, an EHR system will send a request to KanTa.
2) Pharmacies use the ePrescription services provided by KanTa. Their systems access prescription data when making queries about prescriptions and when delivering drugs to patients.
3) Citizens use the Web-browser access provided by KanTa to view their personal health records and prescriptions in the centralised archive.

2.2.2 Service providers (core)

The KanTa services are the core of the architecture. The messaging interface is based on HL7 v3 messaging and used to transport the HL7 CDA R2 documents. The security of the received messages is controlled by the authentication and access control layer. Authentication and digitally signed XML documents are based on the public key infrastructure (PKI) using smartcards, and communication between servers is protected by secure sockets layer (SSL) connections.[15] After a received message has been validated in the messaging interface and the security layer, information is stored in either the eArchive or ePrescription system, depending on the received document type. eArchive and ePrescription are both custom applications implemented by Fujitsu Services Oy on top of the Documentum enterprise content management technology of the company EMC. KanTa utilises a code system so as to have a uniform terminology based on national healthcare classifications and avoid
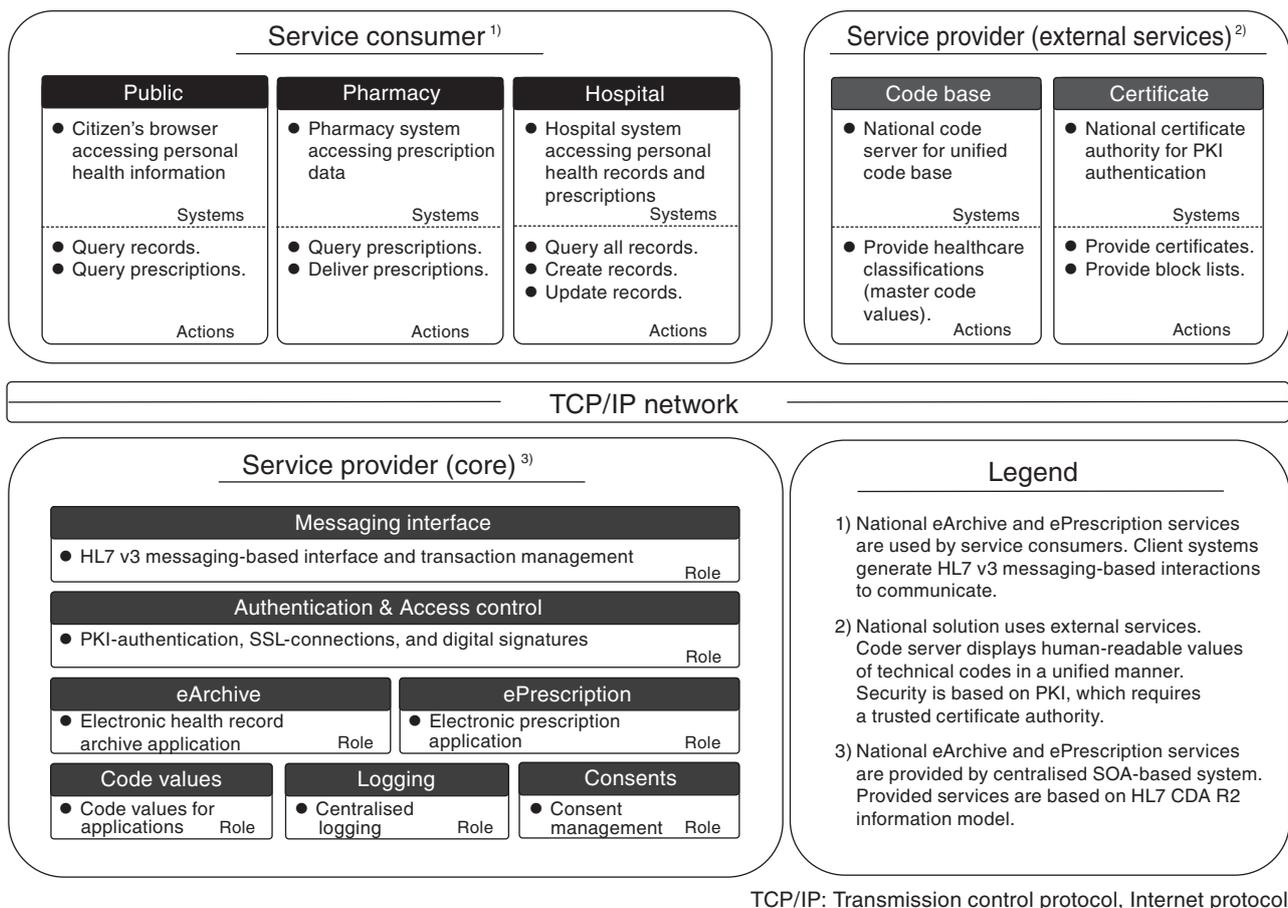
Figure 1
General concept of KanTa services and systems.

typing errors in the field values. The correct code values are verified using the code value service of the KanTa system. All messages are logged in the centralised logging service, which provides detailed information about creating, changing, and reading events in the KanTa system. In Finland, patient health information can be accessed only with the patient's consent, so a service for managing patient consents is an integral part of the KanTa system.[11]

2.2.3 Service providers (external)

External service providers are background services for KanTa, which integrates them and uses provided information to support security by validating certificates using external certificate authority[16] and to synchronise nationally managed healthcare classifications (code values) to KanTa. External service providers play an important role in the KanTa architecture as they provide independent national authority to validate and control information stored in KanTa. For example, the national certificate authority issues personal certificates to healthcare professionals to digitally sign CDA R2 documents. KanTa's authentication and access control layer validates the signatures by utilising the external service provided by the national certificate authority, which also maintains a block list of prohibited certificates. As a result, access to KanTa by healthcare professionals is also managed at a national level.

FUJITSU Sci. Tech. J., Vol. 47, No. 1 (January 2011)

53

3. General information model of KanTa

3.1 General challenge of integration

The ability to manage information in information systems is a major competitive advance in various fields in the public and private sectors. The ability to take comprehensive advantage of available information—for example, to integrate available information sources and assemble relevant information to support decision-making—is one of the key assets of a modern organisation. Moreover, advanced information management practices often play a significant role in supporting organisations at the strategic and operational levels. While decision-making can be strongly supported by integrated information management systems, it is increasingly important to focus on the quality of information. Therefore, from the information quality viewpoint, what is important is not only system integration for exchanging data, but also the ability to build integrations based on the meaning of information.

Integrations focusing purely on data exchange are straightforward to implement. Consider, for example, system *A* that manages client contact information and system *B* that manages a catalogue of services provided by a healthcare organisation. Each system has its own user interface and database solution based on its specific requirements, so system *A* has input fields in its user interface for client address and phone number and a corresponding structure in the database. As system *B* is designed similarly for service catalogue management, the business logic and data management of both systems are based on system-specific rules. However, if a service manager using system *B* would like to access client information stored in system *A*, some integration is needed. The most straightforward way to implement the integration is to directly access the database of system *A* from system *B* and generate user interface fields for system *B* to show the data transferred through the integration. As a result, system *B* reads the data from system *A* and decides how to show it in the user interface of system *B*, or, in other words, determines the meaning of the transferred data as it binds the data to a context in the user interface of system *B*.

On the enterprise scale, the example above is rather problematic. First, integration is entirely based on trusting that system *A* will inform system *B* if there are changes in system *A*. If system *B* does not know about changes, it can bind integrated data from system *A* to a wrong context without knowing it. In an enterprise, there can be tens or hundreds of systems to inform, so change management can be very complex and expensive. Second, a typical integration architecture based on point-to-point integration formulates a complex network where a change in the information model of one system can generate massive modifications to other systems. Finally, there is no common way to communicate between systems. Most integrations are different and data mappings between two systems must be designed individually. In short, there is no semantic compatibility between the information models of different systems.

Typically, mainly for historic reasons, the information possessed by organisations is scattered among various information management systems. These systems are often based on proprietary software implementations and technologies, so stored information is usually semantically incompatible. Semantic noncompliance is one of the main reasons for implementing point-to-point integrations because agreeing on semantic compatibility can be significantly time consuming and difficult. For example, two systems can have the field *Customer*, but in one it refers to a customer inside the organisation and in other to a customer outside the organisation, so it is a lot easier to solve semantic incompatibility between two systems than in the organisation as whole.

In general, the information model and

information architecture can be simplified to make information easier to access or share between various systems. However, as an information model defines information that can be stored in the system, it requires distinct expertise to define and implement an organisation-wide semantically compatible information model that meets the needs of several information systems designed for different tasks. For example, an invoicing application requires data fields for *Purchased Items*, while an application for a service catalogue has very different needs. As a result, a general information model for greatly dissimilar needs can lead to major compromises with all applications within the organisation, and, on the other hand, dissimilar information models between mainly similar applications imply complex point-to-point integrations and poor exploitation of available information. That is, designing and implementing an integration architecture is mostly about understanding the information models used in the various systems across the organisation.

3.2 Principles of KanTa information model

It is a challenge to achieve semantic compatibility between EHR systems on the national scale. The problems arise from organisational aspects as well as for technical reasons. However, when health information archiving is centralised, as it is in Finland's KanTa, there is a path for reaching nationwide semantic compatibility of health information. Semantic compatibility of information stored in KanTa is based on the HL7 CDA R2 standard. CDA R2 provides a standard for describing clinical documents in a structured format.[17] All the elements in a structure are well defined and documented. For example, elements for the patient's name are described in the documentation, and all the integrated systems know the exact location and meaning of the element in KanTa's information model.[9] Therefore, KanTa creates semantic interoperability between the integrated

systems, while all the systems accessing the data must follow the same common information model for integration.

Since HL7 CDA R2 is an international standard, there have been numerous modifications to adjust it to the semantics used in Finnish healthcare. In KanTa, the information model has been updated several times during the implementation project.[9] It has also been found that developing a semantically interoperable information model to be used as the common foundation for integration requires a significant amount of work because changing it requires broad expertise in several areas of information technology, healthcare, and medicine.

The HL7 CDA R2 standard provides the possibility to structure CDA R2 documents from large narrative blocks or to split narrative blocks into smaller elements and use a highly structured format to construct the documents.[17] In KanTa, documents are highly structured and narrative blocks are used only to describe small elements where free text is needed. The inherent benefit of a highly structured information model is that it enables health information to be exploited for various needs. For example, in KanTa, by using standard XML parsers, applications can show the exact part or element of a patient's health record that is needed for a certain therapy, so information can be context sensitive. Without a highly structured format, the benefits of the nationwide health information archive might be significantly smaller. That is, if information were stored inside large narrative blocks, it could not be easily extracted for specific needs. For example, a specific detail about a treatment inside a large narrative block would need to be searched for manually by a healthcare professional. In short, the selected level of the clinical document structure defines the possibilities for exploiting archived information in the future.

4. Discussion

The decisions to build a centralised archive

FUJITSU Sci. Tech. J., Vol. 47, No. 1 (January 2011)

55

and use a highly structured information model are the cornerstones of the KanTa concept. Based on the HL7 v3 messaging standard, KanTa provides a Web Services interface to integrated systems. The interface implements KanTa's highly structured CDA R2-based information model. Therefore, while the integrated systems must comply with the requirements of the KanTa information model, the archived healthcare information is stored in a semantically interoperable format, and the semantic compatibility between the local healthcare systems is achieved through the centralised archive.

One could debate whether to implement the centralised information model and archive or keep information in the local EHR systems, and one could also debate the right level of structure for the CDA R2 documents. These depend strongly on how health information will be utilised. If it is sufficient to access health records as a whole without any context-related automatic processing, then it is rather uncomplicated to build, for example, a reference register model to point to health records in the local EHR systems and access documents as large narrative blocks. However, if automatic health information processing and context awareness for various needs are required, then semantic compatibility and a highly structured information model are needed.

The highly structured information model enables novel possibilities for the utilisation of healthcare information. First, the archived information could be utilised to track the health of the nation. That is, current clinical practices gather information that, on the national scale, can provide detailed information about the development of the nation's health. Using well-known statistical methods, the government could study individual therapies and focus analysis on certain parameters, like blood pressure, and thus combine the results with other socio-economical datasets. Moreover, the structured information

model also provides new possibilities for personal healthcare. The citizen's portal could enable the development of applications that provide tools for personal follow-up for health and potential disease risks based on archived health information. This could significantly improve the results of preventive healthcare as people would be able to get timely information about their own status.

Moreover, recent achievements in systems biology and computational medicine suggest that new methodologies for improved early disease risk assessment are likely to be available in the near future.[18]–[21] These methodologies could provide and utilise information in KanTa and offer vital tools to improve the quality and reduce the costs of healthcare. There are possibilities for improving the quality of care, but the information can also be used as part of medical, technical, and economics research. Therefore, multidisciplinary research of the KanTa system is needed to utilise the possibilities that the new system creates.

5. Conclusion

This paper described the general concepts of the national healthcare information archive of Finland, known as KanTa, which is one of the first nationwide healthcare information archives in the world based on HL7 v3 CDA R2. The real-world benefits of the archive still remain unknown, but examples from the regional reference register systems in Finnish healthcare have led to high expectations for the KanTa system. Moreover, there will soon be actual user experiences, since ePrescription, the first part of KanTa, is expected to be in production in 2010–2011.

References
1) The consolidated legislation of Finland. The act on specialised health care. (In Finnish). *http://www.finlex.fi/fi/laki/ajantasa/1989/19891062?search%5Btype%5D=pika&search%5Bpika%5D=sairaanhoitopiiri*
2) S. Koskinen et al.: Health in Finland. Ministry of Social Affairs and Health of Finland. 2006. *http://www.ktl.fi/hif/*
3) A. Iivari et al.: Health Roadmap—Finland.

Reports of the Ministry of Social Affairs and Health 2007:15.
http://www.stm.fi/en/publications/ publication/_julkaisu/1056833

4) A. Munoz et al.: Proof-of-concept Design and Development of an EN13606-based Electronic Health Care Record Service. *J Am Med Inform Assoc*, Vol. 14, pp. 118–129 (2007).

5) M. W. Bridges: SOA in Healthcare. *Health Management Technology*, Vol. 28, No. 6, pp. 6–10 (2007).

6) E. Vasilescu et al.: Service Orientated Architecture (SOA) Implications for Large Scale Distributed Health Care Enterprises. Proceedings of the 1st Distributed Diagnosis and Home Healthcare (D2H2) Conference Arlington, Virginia, USA, April 2–4, 2006.

7) J. M: Ferranti et al.: The Clinical Document Architecture and the Continuity of Care Record: A Critical Analysis. *J Am Med Inform Assoc*, Vol. 13, pp. 245–252 (2006).

8) The official web page of KanTa project, National Archive of Electronic Health Records.
https://www.kanta.fi/web/en/

9) HL7 Finland document archive. (In Finnish).
http://virtual.vtt.fi/virtual/hl7/cda/ hl7-index.htm

10) Schedule for ePrescription system. The news archive of the ministry of social affairs and health. (In Finnish) 10.11.2009.
http://www.stm.fi/tiedotteet/tiedote/ view/1438004#fi

11) The architecture definition of the KanTa system. (In Finnish).
https://www.kanta.fi/c/document_library/ get_file?uuid=0ae54bdd-b3f0-4c95-9fa9- 68e8275195c3&groupId=10206

12) M. Maass et al.: Regional Health Care Network; efficient and cost-saving. *Finnish Medical Journal*, Vol. 62, pp. 2673–2678 (2007).
http://www.laakarilehti.fi/e/summary. html?opcode=show/news_id=5063/type=4

13) C. M. MacKenzie et al.: Reference Model for Service Orientated Architecture.
http://www.oasis-open.org/committees/ download.php/19679/soa-rm-cs.pdf

14) WS-I Basic Profile 1.1 specification. 24.8.2004.
http://www.ws-i.org/Profiles/ BasicProfile-1.1-2004-08-24.html

15) The architecture for authentication and digital signature in KanTa. (In Finnish).
https://www.kanta.fi/c/document_library/ get_file?uuid=33bad87a-f54a-4d1e-b2ff- 76811aaf4a71&groupId=10206

16) National Supervisory Authority for Welfare and Health of Finland. Certification authority services.
http://www.valvira.fi/en/ supervision_guidance/ca_services

17) R. H. Dolin et al.: HL7 Clincal Document Architecture, Release 2. *J Am Med Inform Assoc*, Vol. 1, pp. 30–39 (2006).

18) P. Soininen et al.: High-throughput serum NMR metabonomics for cost-effective holistic studies on systemic metabolism. *Analyst*, Vol. 134, pp. 1781–1785 (2005).

19) M. Ala-Korpela: Potential role of body fluid 1H NMR metabonomics as a prognostic and diagnostic tool. *Expert Rev Mol Diagn*, Vol. 7, pp. 761–773 (2007).

20) E. Holmes et al.: Metabolic phenotyping in health and disease. *Cell*. Vol. 134, pp. 714–717 (2008).

21) J. K. Nicholson et al.: Systems biology: Metabonomics. *Nature*, *2008;455:1054-6*.

Teemu Suna
Fujitsu Services Oy
Mr. Suna received an M.Sc. degree in Industrial Economics and Management from the Lappeenranta University of Technology, Finland in 2004. In 2007, he joined Fujitsu Services Oy, Helsinki, Finland, where he now works as a Principal Solutions Architect in the Information Management and eServices Business Unit. He is also currently working on his multidisciplinary Ph.D. thesis on information modelling and computational medicine at the Institute of Clinical Medicine of the University of Oulu.

FUJITSU Sci. Tech. J., Vol. 47, No. 1 (January 2011)

57