TECHNICAL REPORT

SEMICOLON: STATE OF THE ART

REPORT NO 2008-0996

DET NORSKE VERITAS

TECHNICAL REPORT

Date of first issue: 1. December 2008	Project No:	DET NORSKE VERITAS AS 1322 Høvik Norway Tel: Fax: http://www.dnv.com NO 945 748 931 MVA
Approved by: Jon Ølnes Principal Engineer	Organisational unit:	
Client:	Client ref.:	

Summary:

Report No: 2008-0996	Subject Group:
Report title: Semicolon: State Of The Art	
Work carried out by: Terje Grimstad, Per Myrseth, Henrik Smith-Meyer, Lasse Udjus, Hans Solli-Sæther, Dag Belsnes, Tore Christiansen, Riitta Hellman, Birger Møller-Pedersen, Audun Stolpe, Arild Waaler, Jim Yang.	
Work verified by:	

Date of this revision:	Revision No:	Number of pages:
	<rev>	143

Indexing terms

Keywords	Service Area
	Market Sector

- ☒ Unrestricted distribution (internal and external)
- ☐ Unrestricted distribution within DNV
- ☐ Limited distribution within DNV after 3 years
- ☐ No distribution (confidential)

Report No: 2008-0996

TECHNICAL REPORT

Table of Contents *Page*

Report No: 2008-0996

TECHNICAL REPORT

Part I

General introduction

TECHNICAL REPORT

1 EXECUTIVE SUMMARY

The Semicolon project aims to improve interoperability in public sector in Norway. The interoperability situation in Norway is not bad; Norway is ranked number 3 in a recent UN survey about e-government readiness which focuses on the evolvement of E-Government; from traditional government via E-Government to "Connected government". However, Norwegian public ICT policy which is defined by the parliament, has defined higher ambitions than actually are met. The Auditor General states that the overall situation for electronic interoperability between public agencies in Norway is not satisfactory compared to expectations and plans. One of the main observations and claims is that electronic information exchange between public organisations would create better and more useful services for private citizens and businesses. Many public organisations are in possession of information which is of value to others, but this information is to a little degree published and made available. Reuse of information would contribute to a more efficient and effective public sector

There are several initiatives and organisations both in Norway and globally which addresses interoperability issues of interest for the Semicolon-project. Some of these are presented in this state-of-the-art report. Part I takes a general look, part II concentrates on organisational aspects and part III on semantic aspects.

The Semicolon-project is comprised of 5 influential public organisations which see large benefits to contributing to a more interoperable public sector. Benefits will be both for the individual organisations it selves, for the public sector and the society as a whole.

Different EU initiatives and organisations are of great interest for Semicolon. IDABC and the research framework programs have contributed a lot to competence and network building in Europe. IDABC's European Interoperability Framework serves as a reference model for interoperability, and its not yet published version 2.0 is further developing the well received concepts from version 1.0. Other initiatives of interest for Semicolon, both finished an ongoing, are project from the EU Information Society Tecghnologies (IST) programme like Athena, SemanticGov,Genesis and Fusion. Interop-VLab serves as a source for information about European projects on interoperability.

In Norway several Ministries and other organizations are working with interoperability issues. Some solutions are in place, but according to the the Office of the Auditor General in Norway the overall situation is not satisfactory compared to expectations and plans.

2 INTRODUCTION

2.1 Norwegian public ICT policy

The interoperability situation in Norway is not bad. The Norwegian public sector is increasingly user-oriented and there is strong national and international policy support for electronic collaboration. 'An Information Society for All', Report no. 17 (2006-2007) to the Norwegian parliament [Report no 17, 2006], addresses electronic collaboration as a means for the provision of electronic services on a 24/7 basis. EU's Lisbon strategy claims that a well-functioning public sector represents a competitive advantage for businesses. It is claimed that collaboration between public organisations, citizens and businesses is necessary to provide a more effective and efficient public sector.

In summer 2007, the Norwegian Ministry of Government Administration and Reform, which is responsible for coordination of the use of information technology and measures to make government more efficient and service-oriented, organised a working group with key ICT-personnel from all major

public organisations to give recommendations about a Common ICT-Architecture for the public sector in Norway. This work was a direct follow up of actions suggested in Report no.17 to the Parliament. The group delivered their report to the ministry in January 2008 [FAOS-report, 2008]. FAOS defines some architecture principles and interoperability is the one which is considered to have largest effect for the improvement of public sector. A common metadataregister is suggested as one of the components in the ICT-Architecture. All the Semicolon-partners from the public sector participated in this working group. The FAOS-report has been out for hearing and the Ministry is now evaluating the different responses.

Quite a lot of infrastructure is already in place. Altinn (www.altinn.no/en) is a service through which citizens and businesses can report information to public authorities. MyPage (www.norge.no/minside) is a portal through which services from different public bodies are made available to the citizens. SERES II is a project run by the Brønnøysund Register Centre and the goal is to provide a national metadata register.

Much of the existing infrastructure and policy developments and requirements for ICT in public sector are relevant for the Semicolon-project. During the years several strategies have been formulated in different policy documents.

For the health sector there is a series of strategy documents which focuses on the use of electronic services to provide better collaboration and communication and thus a better health service.

- ICT for a better health service, action plan 1997-2000 [Health plan, 1997]
- Electronic collaboration in the health and social sector, action plan 2001-2003 [Health plan, 2001]
- Te@mwork 2007, Electronic Cooperation in the Health and Social sector, National strategy 2004-2007 [Health plan, 2004]
- Teamwork 2.0, National strategy for electronic collaboration in the health and social sector 2008-2013 [Health plan, 2008]

The ICT policy of the government is also described in a series of documents.

- Strategy for ICT in the Public Sector (2003-2005) [eGov plan, 2003]
- eNorway 2005 [eGov plan, 2005]
- eNorway 2009, The Digital Leap [eGov plan, 2009]
- Report no 17 to the Storting (2006-2007), An information society for all [Report no 17, 2006]

Even though there has been policy support for interoperability during the last two decades, and several services are in place, many unresolved issues still remain.

One of the conclusions from the FAOS-group was that the stove-piped managerial structure in the public sector is an obstacle for collaboration and communication. The public sector organisations are allocated a budget by their ministry. All of the goals and evaluation criteria are intra-organisational. No, or only very limited, funds are allocated for collaboration activities spanning several organisations belonging to different ministries.

The Office of the Auditor General in Norway has made a recent study of ICT as an instrument to obtain a better health service and better utilisation of resources in the Health sector in Norway [Riksrev

Doc 3:7, 2008]. The study concludes that still, after 10 years with heavy investments in ICT-support, a vast majority of information exchange is paper-based. The Directorate for Health and Social Affairs has had the responsibility to implement a national strategy plan for electronic collaboration in the health sector. However, it seems that the means available, and the means in this context should not be interpreted as money only, has been too limited (or have not had the desired effect). The Auditor General raises the question whether there is a discrepancy between the responsibility and the means available to reach the target.

Another report from the Auditor General [Riksrev Doc 3:12, 2008] raises the same questions for electronic interoperability in general between public agencies in Norway, stating that the overall situation is not satisfactory compared to expectations and plans. One of the main observations and claims is that electronic information exchange between public organisations would create better and more useful services for private citizens and businesses. Many public organisations are in possession of information which is of value to others, but this information is to a little degree published and made available. Reuse of information would contribute to a more efficient and effective public sector.

Most of the services available today stem from one public body only. There are almost no cross sector services, e.g. services to citizens in life-cycle situations where a child is born or a person dies. Such services require streamlined business processes running through several public organisations. There must be integration with existing data registers, and a common view of the information following the processes e.g. represented and supported by a national metadata register. Last, but not least, an easy to comprehend, universally designed user interface to the services through well known public portals must be supported. Few of these assets are in place.

The aim of the Semicolon-project is to provide methods, tools, metrics and competence to contribute to a more interoperable public sector.

2.2 Research directions in Semicolon

This report identifies a set of research directions of interest to the Semicolon project. This report is limited to (i) the scope and goals of the project and (ii) by the project team's collective knowledge and the literature we have found and chosen to use.

The first limitation is the Semicolon scope and goal: *Develop and test ICT-based methods, tools and metrics, to obtain faster and cheaper semantic and organisational interoperability both with and within the public sector.*

The building blocs of this goal are as follows:

1. within the domain of public sector
2. obtain organisational interoperability
3. obtain semantic interoperability
4. obtain 1 & 2 faster and cheaper
5. obtain 1,2 &3 by develop and test ICT-based methods, tools and metrics

The second limitation is related to the collective project knowledge and the literature we have chosen to use.

There is a growing science base within Enterprise Interoperability (EI) of relevance to interoperability [Cordis EI Roadmap, 2006]. An aspect of the EI Science base considers e.g. how to use System Theory to model EI [INTEROP-VLab] .

The interoperability issues related to collaborating organisations can range issues related internal to an enterprise, between enterprises (cross government) and further to pan-European services and interoperability.

The relation between Semicolon research challenges, Semicolon results and our science base is illustrated in Figure 1.

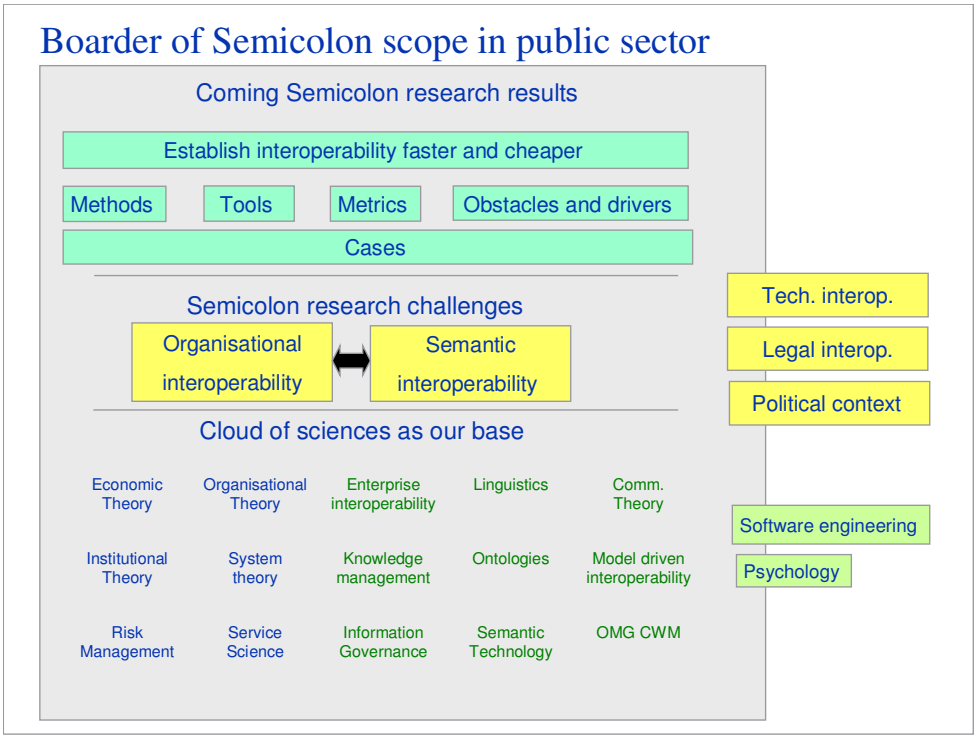


Figure 1, Semicolon scope in public sector

The breakdown of this report is therefore based on the areas of challenge, and under each such area there is a breakdown based on a puzzle of subject from our science base.

3 ABOUT THE SEMICOLON PROJECT

Semicolon (Semantic and Organisational Interoperability in Communicating and Collaborating Organisations) is a three year R&D- project partly funded by the Norwegian Research Council. http://www.semicolon.no/semicolon-web/Hjemmeside-E.html.

The main goal of Semicolon is to develop and test ICT-based methods, tools and metrics to obtain faster and cheaper semantic and organisational interoperability both with and within the public sector.

Semicolon's sub-goals are to:

- Identify obstacles for interoperability and strategy/solutions to tackle these.

- Develop methods and tools to establish effective and efficient interoperability.
- Develop metrics and indicators to measure the effects of interoperability (including use of ontologies).
- Apply methods and tools to develop ontologies and models for use in concrete cases in public organisations.
- Measure effects by the use of the metric.
- Disseminate results and experiences to public sector, ICT industry and academia.
- Create an arena with influential public organisations, competent companies and research networks, as a locomotive for semantic and organisational interoperability in public sector.

Five large and influential public bodies are engaged in Semicolon. These will provide real world collaboration example cases as study items for the project. The participants of Semicolon are: the Directorate of Taxes, the Brønnøysund Register Centre, the Directorate for Health, Statistics Norway and The Norwegian Association of Local and Regional Authorities – KS. The organisations performing the research are Det Norske Veritas (DNV, the coordinator and project owner), the company Karde (initiator of the project), consultancy Ekor and the Norwegian Centre for Informatics in Health and Social Care - KITH. The University of Oslo and the Norwegian School of Management as well as the universities of Oxford and Aberdeen provide expertise in semantics, object orientation and organisational theory.

The SEMICOLON-consortium is multidisciplinary. It aims at identification of obstacles in real collaboration cases, and at the development of new methodologies, tools and metrics. The development work shall be based on prototypes, and the results shall be verified in real collaboration cases. SEMICOLON allows differences in collaborating organisations, such as different values, different goals, and different aspects of the same concept.
.

4 CHARACHTERISTICS OF NORWEGIAN PUBLIC SECTOR

This State of the art is concerned with interoperability among about 700 autonomous governmental bodies and the services that they offer to business, citizens and visitors. Obligations and rights are given by law at various levels of localization (Human Rights, EU, and National) and by a practice that is expected to be fair and equal-for-all.

When attempting to understand the challenge of collaboration and interoperability within this domain, it is important to understand some of the domain characteristics and how it might differ from other domains (such as Collaborative Engineering), - in order to better evaluate methods, tools and techniques in other domains.

There are about 700 governmental bodies, each with their tasks and obligations given by law. Of these, there are 429 Municipalities with local responsibilities, but with common representation in The Norwegian Association of Local and Regional Authorities. In other words, it is “by nature” highly federated. Change might happen slower than in other domains, since a change of law may be a prerequisite for the change. The process of changing the law is by intention slow, giving democratic right for everyone to be heard and making sure the need for change is real and not “just the latest fashion or a fix idea.”

There is no central, governmental body for operation and development of ICT-based methods and solutions. Each of the 700 bodies are responsible for their own way of doing things, only to some

degree guided by standardization, - but with no or little operational support, or a common operational platform to build upon and utilize. A number of central initiatives and sector portals exist (www.altinn.no , www.minside.no, www.regelhjelp.no, www.brreg.no, www.norge.no , www.regjeringen.no to mention some). However, none of these offer a common platform for the bodies to utilize *themselves*, thereby ensuring consistency to form and content structure (user interfaces, logical and conceptual breakdown, navigation, etc.). Two promising seeds are MinSide, from which authenticated governmental services can be offered and reached, and AltInn, developed by the "leading Governmental body" of Taxation. The next version even promises business people to build and support interactions without the need of software programming, but it is still far from guaranteed that others will be able and willing to "join" and utilize the platform. Instead, each body publishes their own information, - on their own form and place. If one wanted to write algorithms to analyze the available information, thereby evaluating quality and coverage associated with some criteria or measure, one would have an unnecessary hard challenge.

Most of the governmental bodies have their own IT department, - responsible for IT systems and support. The larger groups include a couple of hundred people, with considerable use of consultants. Centralization of service "might not be" in the interest of those providing the services to day. Furthermore, the ability to come up with a quick fix to a local problem there and then might give better feedback than what it is possible to achieve by trying to cooperate with the other 700 bodies. It might also be quicker and cheaper to start from scratch and realize local solutions than to spend time to understand what others have been doing and to utilize other peoples work. Even if it might be obvious that such distributed organization is inefficient and probably even counterproductive, a quick reorganization is not to be expected in the near future. Even if it might be obvious that such organizations are inefficient and probably even counterproductive, a quick reorganization is not to be expected in the near future. Even with the obvious change of technological platform from the distributed client-server to the web that "makes the world a smaller place", where everybody can share the same operational platform, and where the centralizing drivers are becoming stronger and stronger, - human factors of established power, position, operational dependency and cost of changing current systems that were not designed for change, - will work to maintain status quo and resist the systems that this sector needs.

They have some freedom in how they exercise their obligations, in particular, the municipalities have the same tasks and obligations, but are expected to prioritize and adapt their realization to local conditions and utility.

The information model ("product model") is "shallow", with some obvious potential Master Data (Peoples name, address, bank account) being relevant in numerous context, - but without national standardization.

It is possible and not uncommon to challenge any decision or product provided by a governmental body (a right,that is also commonly executed to the extent that is considered common practice in some domains). Some decisions can even be challenged at multiple levels. Many decisions are complex, taken in isolation from similar or related decisions, and an objective truth is hard or impossible to achieve. Even with attempts to regulate and standardize, there will be subjective opinions and evaluations that must be open to challenge. This also implies, contrary to the process of building a spacecraft or running a power plant, that the need and possibility to "do it right the first time" is far less important.

The governmental services are intended to be fair and equal, making sure that the strong and powerful are treated equally with the actors who are perhaps not aware of their rights, obligations and

possibilities (of for example applying for available funds or services in given situations, - without any right of receiving anything), or who are not able to execute whatever is expected of them in order to achieve their maximum utility.

A considerable part of the information is of personal nature and must only be available to those who can document a legitimate need. This imposes challenges to the open exchange of information on the internet, - and even on potential "governmental networks" where far from all of the 700 governmental bodies are eligible to see any given part of information available. Knowing and supporting who, - in what role and in context of what objective and responsibility, has sufficient authentication

- to *know*/learn about the *existence* of any given information, (still not learning the content itself)
- to gain access to the information,
- to update/change the information,
- to delete the information, and
- to change the information about the requirements of authentication and authorization.

To some extent, we can say that the complexity of security and sensitivity is far greater in the Public Sector than in other domains, while the information itself is far simpler.

5 NORWEGIAN ORGANISATIONS

There are several organisations in Norway addressing tasks related to interoperability.

5.1 Public policy

The public policy represents an ambition for much interoperability in public sector. As discussed earlier, the situation is not bad, but the Auditor General concludes that much work remains to implement the rather high expectations.

5.2 Ministry of Government Administration and Reform (FAD)

Source: http://www.regjeringen.no/nb/dep/fad
The Ministry of Government Administration and Reform is responsible for the ICT policy in Norway.
Relevance for Semicolon
Interoperability is very much a policy issue, and implementation of Semicolon results and suggestions needs to be endorsed by FAD.

5.3 Ministry of Justice and the Police (JD)

Source: http://www.regjeringen.no/nb/dep/jd
The Ministry of Justice has made an information model for criminal cases, from report a person to the police to sentence. The Ministry is about to define a metadataregister of concepts which should be used for legislation.
Relevance for Semicolon
The information model can give input to Semicolon, and the metadata work may exploit Semicolon results.

5.4 Ministry of Trade and Industry (NHD)

Source: http://www.regjeringen.no/nb/dep/nhd

One of the responsibilities of the Minstry of Trade and Industry is simplification of public reporting from businesses. Altinn and metadata strategies are elements of the simplification work. The Brønnøysund Register Centre is overseen by NHD.

Relevance for Semicolon

Altinn, metadata and simplification are important instruments and topics for interoperability and as such of interest for Semicolon.

5.5 Statistics Norway (SSB)

Source: www.ssb.no

Statistics Norway has defined a metadata strategy which was approved by top management early 2005. SSB has launched a metadata portal (www.ssb.no/metadata). They have also initiated a work on an enterprise model.

Relevance for Semicolon

SSB is a partner in Semicolon. If organisations should be able to interoperate, it is necessary to know and have control over their own processes, systems and information. Furthermore, you need to know the other partners processes and information. In other words you need to have some kind of enterprise model and metadata repository. SSB experiences are important to incorporate in Semicolon work.

5.6 Tax Authorities

Source: www.skatteetaten.no

The public sector is the driving force when applying ICT for services, with the Tax Directorate (SKD) as the locomotive. A public web-based self-declaration service was launched in 1999. At the same time, the Directorate enabled reporting for companies via the web and directly from enterprise resource planning (ERP) systems. Tax is the driving force behind the Altinn-initiative. Now they have launched the eDialogue-initiative. eDialogue services are cross sector services such as the eDialogue related to birth ("birth dialogue") and eDialogue related to death ("death dialogue"). The eDialogue requires modelling of cross sector processes and accompanying information.

Relevance for Semicolon

SKD is a partner in Semicolon. They have high implementation power. eDialogue is one of the cases in Semicolon. The success of Semicolon is dependant on active collaboration with the SKD.

5.7 Health sector

Source: www.kith.no and www.shdir.no

KITH (Norwegian Centre for Informatics in Health and Social Care) established and maintains a metadatabase for the Norwegian healthcare sector. The content of the metadatabase is made available through the website www.volven.no. The metadatabase consists of (intentionally) various types of metadata such as coding and classification systems (for coding/classifying health-related as well as organisational/technical information), concept definitions, data/technical standards such as messaging standards.

Relevance for Semicolon

KITH and Norwegian Directorate of Health (in Norwegian Helsedirektoratet) representing the health sector are partners in Semicolon. In the same way as argued for SSB above, results and experiences from the health sector are important to incorporate in Semicolon project.

5.8 Brønnøysund Registre Centre (BR)

Source: www.brreg.no

The Brønnøysund Register Centre is operating a lot of common public registers and is operating the Altinn service. BR is also running the SERES II-project which aims to create a national metadata register.

5.9 Norwegian Mapping Authority (Statens kartverk – SK)

Source: www.statkart.no
SK is responsible for Norway Digital. Norway Digital is public infrastructure for geographic information for the entire society. The objectives for Norway Digital are:

- More efficient private sector e.g. within transportation at land, sea and in air
- More efficient public sector, better service to the public, better decision-making and increased participation
- Better management of areas environmental and natural resources
- Increased safety and readiness

There are many areas of applications for geographic information: municipal management; land transportation (route planning and fleet administration); shipping and fisheries (navigation); agriculture and forestry (planning); defence and crisis handling; telecomm, energy- and road administration; environmental and resource management; oil protection; property management; media; recreation; map production.

Relevance for Semicolon
SK interoperability experiences is a valuable repository and important to take into account for Semicolon activities, both on the organisational and semantic interoperability level.

5.10 DIFI - Agency for Public Management and eGovernment

Source: www.difi.no
The Agency for Public Management and eGovernment (DIFI) was established 1 January 2008, following a merger of the previous public agencies Statskonsult, Norway.no and the Norwegian eProcurement Secretariat.

The Agency for Public Management and eGovernment (DIFI) aims to strengthen the government's work in renewing the Norwegian public sector and improve the organisation and efficiency of government administration. DIFI works to ensure that government administration in Norway is characterised by values of excellence, efficiency, user-orientation, transparency and democracy.

DIFI's priorities in the ICT area are:

1. Secure eGovernment
2. Standardisation and architecture
3. Accessibility of services through MinSide and Norge.no
4. Management and coordination
5. International activities

Source: Speach by Director of DIFI, Hans Christian Holthe, "Development through cooperation", NOKIOS 2008 conference in Trondheim, 16th October 2008.

The agency is overseen by the Ministry of Government Administration and Reform (FAD) which is responsible for the ICT policy in public sector. DIFI is FAD's instrument to implement the ICT-policy.

Relevance for Semicolon
DIFI is foreseen as a collaborating partner for both utilising and disseminating Semicolon results.

5.11 NorStella - Foundation for e-Business and Trade Procedures

Source: www.norstella.no
Norstella was established 1 January 2003 and is appointed by the Norwegian Government as the national contact point for all international standardization activities in the field of electronic business and trade facilitation. It is a user oriented, independent, non-profit, private foundation located in Oslo. NorStella deals with Internet-based standards like ebXML, Core Components, UBL and semantic technologies. And it covers all kinds of industry.

The main objective of NorStella is to contribute to effectiveness and efficiency in public and private undertakings, by promoting

- rational and simplified processes and procedures in national and international trade, in the private as well as the public sector and between private companies and public administrations
- reliable and efficient implementation and profitable use of standardized data interchange between organizations

Two Norstella-committees are especially relevant in the interoperability context.

Committee on Interoperability
The objective of the InterOp-committee is to give advice, quality insurance, recommendation and use of international standards for semantic interoperability.

Committee on Trade Facilitation
The objective of the Trade Facilitation committee is to facilitate and disseminate common standards for international trade.

Relevance for Semicolon
Norstella helps to obtain one of the sub goals of Semicolon: To create an arena with influential public organisations, competent companies and research networks, as a locomotive for semantic and organisational interoperability in public sector.

Semicolon is an active partner in the InterOp-committee and Sermicolon and Norstella organised a rather successful common seminar 5th November with the topic Electronic collaboration in Public Sector.

Semicolon and Norstella are Norwegian partners in the Semic.eu initiative.

6 INTERNATIONAL ORGANISATIONS AND INITIATIVES

There are several organisations and initiatives which are of interest and relevance for Semicolon.

6.1 IDABC EIF2.0 Draft - European Interoperability Framework

Source: http://ec.europa.eu/idabc/

TECHNICAL REPORT

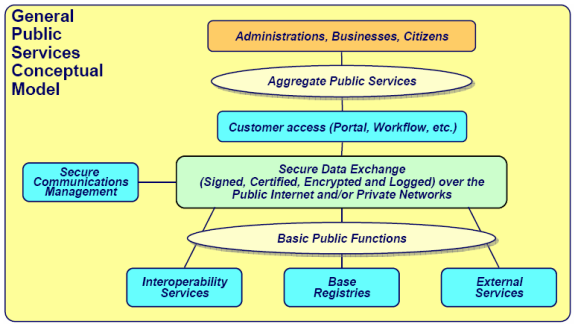
IDABC (Interoperable Delivery of European eGovernment Services to public Administrations, Businesses and Citizens) is a European Union Program that promotes the correct use of Information and Communication Technologies (ICT) for cross-border services in Europe.

The European Interoperability Framework of the IDABC is concerned with the interoperability between Member States of the European Union. Important elements of the EIF initiative include the EIF specifications themselves, the Interoperable Delivery of European eGovernment Services (IDABC) program, a set of Architecture Guidelines (AG), and a program for Pan European Government Services (PEGS) based on a Generic Public Services Conceptual Model (GPSCM). EIF version 1.0 was well received upon publication in 2004. It "provides recommendations and defines generic standards with regard to organizational, semantic and technical aspects of interoperability, offering a comprehensive set of principles for European cooperation in eGovernment."

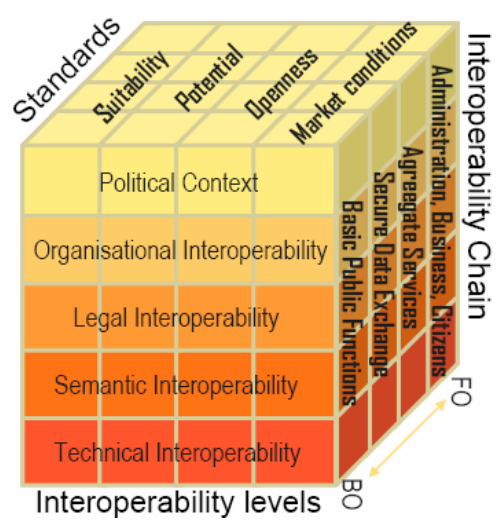
Related issues and patterns can be repeated at the national level below, between Governmental Bodies within a member state. In general a hierarchical pattern of localization is perceivable.

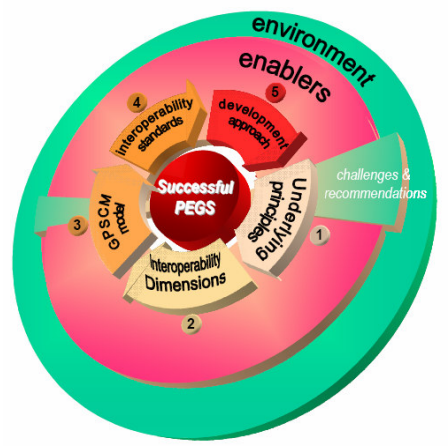
The EIF is currently under revision towards Version 2.0 [IDABC EIF 2.0]. A Draft document has received comments from the community and is expected to be published early 2009. In the draft, we find the following proposals:

1. A Political Context and a Legal Interoperability Level in addition to the Organizational, Semantic and Technical interoperability levels of the EIF1.0.
2. A proposal for Pan-European Government Services - PEGS
3. A proposal for a Generic Public Services Conceptual Model (GPSCM), as illustrated in the figure
4. Guidelines on development approach for the EU and National Frameworks that wish to comply

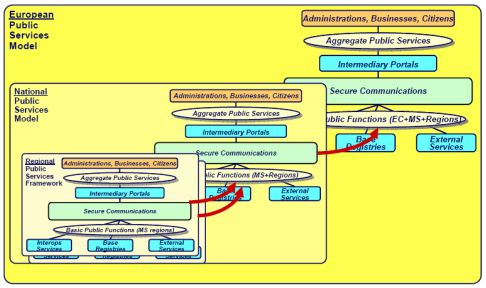


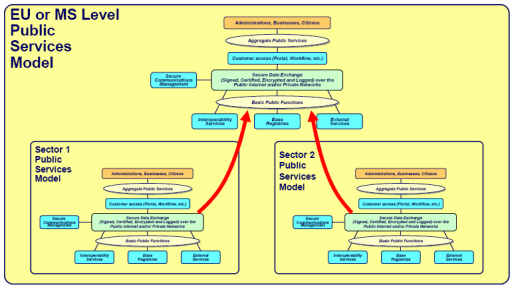
The expressed goal of the EIF2.0 Draft is "to support the development and deployment of PEGS at the conceptual level". The following figures illustrate the three dimensions of the EIF2.0 Draft together with an overview of the EIF elements in an environment of enablers:





The draft addresses both the "cross-border" case and the cross-sector case within the framework, as illustrated in the following figures:





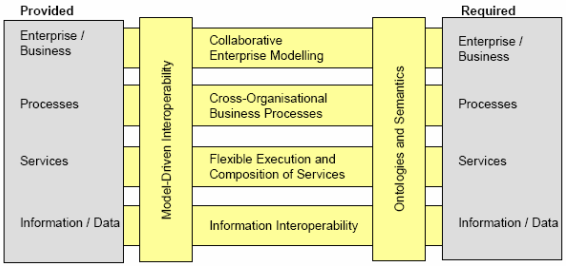
Relevance for Semicolon

The framework is relevant to Semicolon as a reference for communication and relation to other relevant development and work. Even if the EIF is concerned with interoperability between Member States and depend on national frameworks to address National interoperability, the *patterns* of challenges and *patterns* of possible solutions are, - or should be repeatable at different levels. By utilizing similar patterns of solutions on multiple levels, we build collective confidence, recognition and abilities to provide solutions that can be shared, understood and repeated wherever useful and possible, - enabling interoperability.

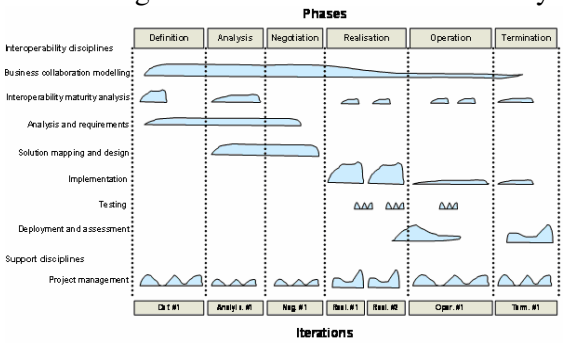
6.2 AIF - Athena Interoperability Framework

The Athena Interoperability Framework (http://modelbased.net/aif/framework.html) is a result of the Athena-project (described in chapter 10.4.1 in part II below), covering for example

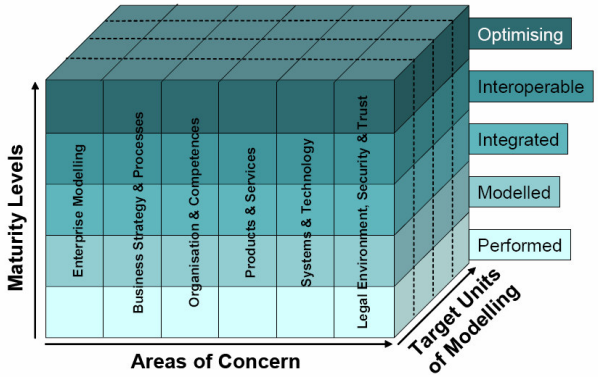
- An Interoperability Reference Architecture



- An Interoperability Methodology Framework, where the applicability of the various methodologies also is set in context of life cycle phases.



- An Enterprise Interoperability Maturity Model (EIMM), where the three dimensions are illustrated in the following figure:



- An identification, description and Classification of 29 Interoperability Issues[1] (Challenges) within Business Management, Process Management, Knowledge Management, Information management, Software management, and Data Management

[1]Athena Classification of Interoperability Challenges:
http://modelbased.net/aif/reference_architecture/characterisation.html

Relevance for Semicolon

The framework is relevant to semicolon, with its identification of challenges and multidisciplinary approach where methods are identified in relation to their context of best use. We note however also that the project has its roots and scenarios in collaborative product engineering and that collaborative, distributed modeling on the internet has entered the scene after project completion.

6.2.1 Relations between EIF and AIF

The AIF is also influencing the current IDABC work on an Enterprise Interoperability Framework, which in the EIF1 and a 2007 "version 2" was mapped as follows: (from Athena DA4.2 page 114)

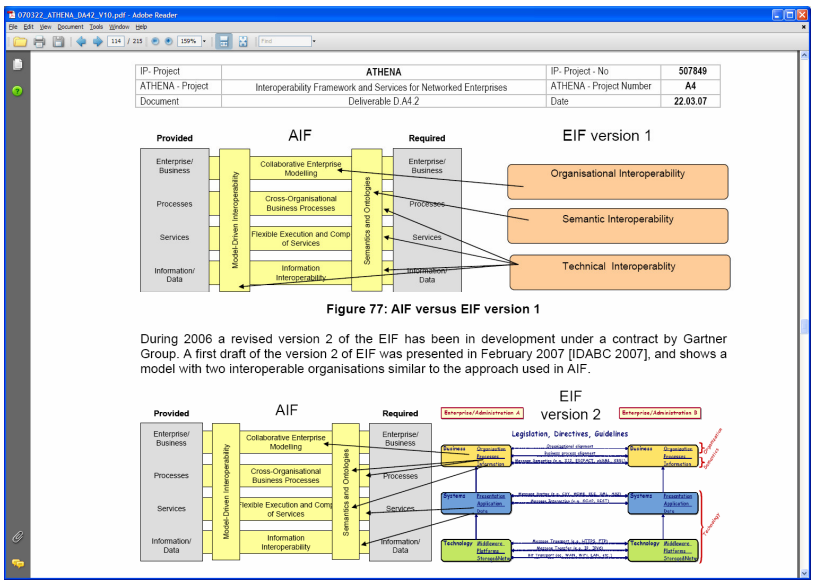


Figure 77: AIF versus EIF version 1

During 2006 a revised version 2 of the EIF has been in development under a contract by Gartner Group. A first draft of the version 2 of EIF was presented in February 2007 [IDABC 2007], and shows a model with two interoperable organisations similar to the approach used in AIF.

See more on IDABC in chapter 6.1 above, where a new version 2 is proposed in the current draft.

6.3 United States Federal Enterprise Architecture

The Federal Enterprise Architecture (FEA) is an initiative of the US Office of Management and Budget that aims to comply with the Clinger-Cohen Act and provide a common methodology for information technology (IT) acquisition in the United States federal government. It is designed to ease sharing of information and resources across federal agencies, reduce costs, and improve citizen services.

The FEA is currently a collection of Reference Models that develop a common taxonomy and ontology for describing IT resources. These include the Performance Reference Model, the Business Reference Model, the Service Component Reference Model, the Data Reference Model and the Technical Reference Model.

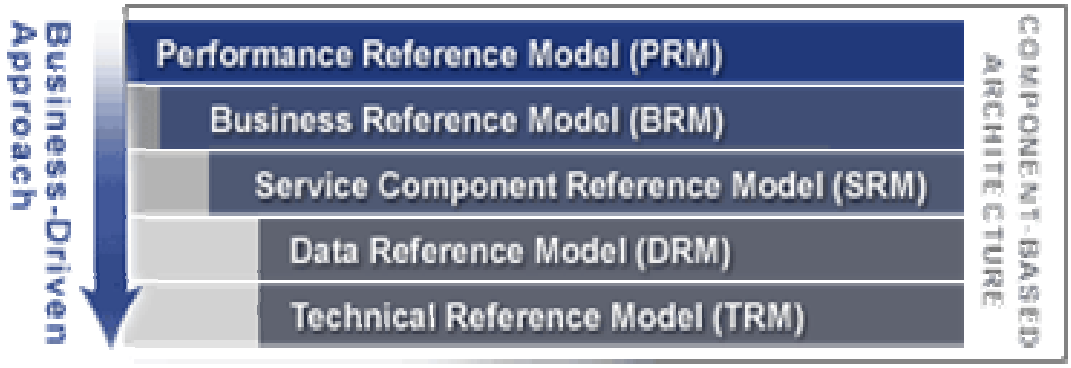


Figure 2: Federal Enterprise Architecture

The **Performance Reference Model (PRM)** is a standardized framework to measure the performance of major IT investments and their contribution to program performance. The

PRM has three main purposes:

- Help produce enhanced performance information to improve strategic and daily decision-making;
- Improve the alignment — and better articulate the contribution of — inputs to outputs and outcomes, thereby creating a clear "line of sight" to desired results; and
- Identify performance improvement opportunities that span traditional organizational structures and boundaries

The PRM uses a number of existing approaches to performance measurement, including the Balanced Scorecard, Baldrige Criteria, Value Measurement Methodology, program logic models, the value chain, and the Theory of Constraints. In addition, the PRM was informed by what agencies are currently measuring through PART assessments, GPRA, Enterprise architecture, and Capital Planning and Investment Control. The PRM is currently composed of four measurement areas:

- Mission and Business Results
- Customer Results
- Processes and Activities
- Technology

The **Business Reference Model (BRM)** is a function-driven framework for describing the business operations of the Federal Government independent of the agencies that perform them.

The Business Reference Model provides an organized, hierarchical construct for describing the day-to-day business operations of the Federal government using a functionally driven approach. The BRM is the first layer of the Federal Enterprise Architecture and it is the main viewpoint for the analysis of data, service components and technology.

The BRM is broken down into four areas:

- Services For Citizens
- Mode of Delivery
- Support Delivery of Services

- Management of Government Resources

The **Service Component Reference Model (SRM)** is a business and performance-driven, functional framework that classifies Service Components with respect to how they support business and/or performance objectives.

The SRM is intended for use to support the discovery of government-wide business and application Service Components in IT investments and assets. The SRM is structured across horizontal and vertical service domains that, independent of the business functions, can provide a leverage-able foundation to support the reuse of applications, application capabilities, components, and business services.

The SRM establishes the following domains:
- Customer Services
- Process Automation Services
- Business Management Services
- Digital Asset Services
- Business Analytical Services
- Back Office Services
- Support Services

The **Data Reference Model (DRM)** describes, at an aggregate level, the data and information that support government program and business line operations. This model enables agencies to describe the types of interaction and exchanges that occur between the Federal Government and citizens.

The DRM categorizes government information into greater levels of detail. It also establishes a classification for Federal data and identifies duplicative data resources. A common data model will streamline information exchange processes within the Federal government and between government and external stakeholders.

Volume One of the DRM provides a high-level overview of the structure, usage, and data-identification constructs. This document:
- Provides an introduction and high-level overview of the contents that will be detailed in Volumes 2-4 of the model;
- Encourages Community of Interest development of the remaining volumes; and
- Provides the basic concepts, strategy, and structure to be used in future development.

The DRM is the starting point from which data architects should develop modeling standards and concepts. The combined volumes of the DRM support data classification and enable horizontal and vertical information sharing.

The current published version of the DRM is undergoing revision. The Project Management Office of the FEA (FEA PMO) is collaborating with members of the interagency DRM working group, chartered by the Architecture and Infrastructure Committee (AIC) of the Chief Information Officer (CIO) Council, to further enhance and improve this reference model.

The **Technical Reference Model (TRM)** is a component-driven, technical framework used to categorize the standards, specifications, and technologies that support and enable the delivery of service components and capabilities.

The Technical Reference Model provides a foundation to categorize the standards, specifications, and technologies to support the construction, delivery, and exchange of business and application components (Service Components) that may be used and leveraged in a Component-Based or Service-Oriented Architecture. The TRM unifies existing Agency TRMs and E-Gov guidance by providing a foundation to advance the re-use of technology and component services from a government-wide perspective.

The Service Areas in the TRM are:
- Service Access and Delivery
- Service Platform and Infrastructure
- Component Framework
- Service Interface and Integration

Together these five Reference Models describe Information Systems and associated enterprises from technical, linguistic, semantic and pragmatic points of view, and can be used as basis for mandatory specification and analysis of all proposed major US ICT development projects and infrastructure investments.

6.4 Interop-VLab – a source for information about European projects on interoperability

Source: http://interop-vlab.eu/interop-vlab-network/INTEROP-VLab/
On their website, the VLab describes themselves as being

- A network of **8 regional poles**, bringing together leading academics, research centers, industrial stakeholders, SMEs, from 9 European countries and from China
- An access route to **200 top specialists** in the domain of **Enterprise Interoperability** (EI)

They propose

- "Solutions which cut across organization, semantic requirements and IT
- Dissemination of Knowledge; Education & Training; Collaborative analysis of the market situation; State of the Art; Standardisation advice; Mutualisation of research funds through Collaborative Research"

Semicolon use this site as useful source for our study of relevant European Projects and Contributions to Interoperability Research, as described in chapter **Feil! Fant ikke referansekilden.**.

6.5 National European programmes

The "European Interoperability Framework" version 1.0 (EIF V1.0) was extremely well received in the world of public administrations in Europe (and elsewhere) and is often referenced as one of the founding documents when interoperability is discussed.

Many Member States of the European Union have launched efforts on interoperability, and have ahead established National Interoperability Frameworks or similar interoperability guidelines in order to provide guidance to project managers and procurement officers.

Belgium
http://www.belgif.be

Denmark
http://standarder.oio.dk/my-home-your-home/view?set_language=en

Estonia
http://www.riso.ee/en/information-policy/interoperability

France
http://synergies.modernisation.gouv.fr/rubrique.php?id_rubrique=1

Germany
http://www.kbst.bund.de/cln_012/nn_837392/SharedDocs/Meldungen/2006/saga____0.html

Ireland
http://www.reach.ie/technology/interoperability.html

Italy
Technical framework: http://www.cnipa.gov.it/site/it-IT/In_primo_piano/Sistema_Pubblico_di_Connettivit%c3%a0_(SPC)/Servizi_di_interoperabilit%c3%a0_evoluta_e_cooperazione_applicativa/
Legal framework: http://www.cnipa.gov.it/site/_files/Opuscolo%2013II.pdf

Malta
http://ictpolicies.gov.mt/docs/cimu_t_0001_2002.pdf

Netherlands
http://www.e-overheid.nl/atlas/referentiearchitectuur/
http://www.e-overheid.nl/data/files/architectuur/E-government_in_the_Netherlands.pdf

Spain
http://www.csi.map.es/csi/pg5c10.htm

Sweden
http://www.verva.se/shs and http://www.verva.se/framework

United Kingdom
http://www.govtalk.gov.uk/schemasstandards/egif_document.asp?docnum=949

The portal to on-line European and national public services http://ec.europa.eu/youreurope/ contains access to a mixture of European and national resources

6.6 Semic.eu

Source: www.semic.eu
Semicolon is a partner with Semic.eu. SEMIC.EU (Semantic Interoperability Centre Europe) is an EU-Project to support the data exchange for pan-European eGovernment services. It aims to create a repository for interoperability assets that can be used by eGovernment projects and their stakeholders.

6.7 SemanticGov

Source: www.semantic-gov.org
SemanticGov (EU funded FP6-2004-IST-4-027517, 2006-2008) aims at building the infrastructure (software, models, services, etc) necessary for enabling the offering of intelligent services by public administration through the use of the semantic web. Through this cutting edge infrastructure, SemanticGov will address longstanding challenges faced by public administrations such as streamlining cooperation (e.g. through achieving interoperability) amongst public administration agencies both within a country as well as amongst countries, easing the discovery of public administration services by its customers, facilitating the execution of complex services often involving multiple PA agencies in interwork-flows.

To achieve this, the SemanticGov project aims at capitalizing on the Service Oriented Architectures paradigm, implemented through state-of-the-art Semantic Web Services technology and supported by rigorous and reusable public administration domain analysis and modelling, while being in line with all major European programmes and initiatives in the field such as the European Interoperability Framework and the recent work conducted by the EU IDABC Programme, the forthcoming i2010 group of Member States representatives and the Competitiveness & Innovation (CIP) Programme. [SemanticGov, Extended Project Presentation] [SemanticGov, Conceptual Analyses]

The deliverables is related to architectures, several types of models and suggestions on protocols and standards. Examples of deliverables are[2]:

- A formal model for a Public Administration service on the basis of the Web Service Modelling Ontology (WSMO).
- Reengineering the public administration modus operandi through the use of reference domain models and Semantic Web Service technologies,
- An Interoperability Framework for Pan-European E-Government Services (PEGS).
- Mapping Citizen Profiles to Public Administration Services Using Ontology Implementations of the Governance Enterprise Architecture (GEA) models.

The suggested SemanticGov infrastructure consists of:
1. The Needs-to-Services facilitator (N2S facilitator)

[2] http://www.semantic-gov.org/ + menu choice publications.

2. The National Public Administration Service Directory (NPASD)
3. The Distributed Business Process Manager (DBPM), as the infrastructure for on-the-fly, semi-automated composition, execution, and monitoring of complex Public Agency Services
4. The Public Agency Ontology Server, as the knowledge infrastructure that manages the PA domain models
5. The Communal Semantic Gateway (CSG), which resolves semantic incompatibilities amongst different public administration systems

The project runs from 2006 to end of -2009.

7 INITIATIVES AND BEST PRACTICES IN NORWAY

7.1 Public sector

All initiatives, services and projects mentioned in this chapter are related to the interests of Semicolon.

7.1.1 The Register of the Reporting Obligations of Enterprises

The Register of the Reporting Obligations of Enterprises maintains an ongoing overview of the reporting obligations of business enterprises and finds ways of simplifying all the paperwork. The aim is to remove superfluous data compilation and information registration, particularly in consideration of small and medium-sized businesses.

The Register of the Reporting Obligations of Enterprises shall indicate what type of information can be found in the various registers and public agencies. Data shall be sent to the individual agencies as before.

A number of acts and regulations require the public authorities to have detailed information about various categories of business enterprises. Many reporting obligations are quite similar for small and large business enterprises, even if the large companies have more resources for paperwork and administration. Therefore small and medium-size companies suffer disproportionately, while public agencies also must allocate many resources to collecting and processing information received from the companies.

In many cases various public agencies ask companies the same or quite similar questions, which often mean unnecessary duplication of work for both companies and the authorities. With no total overview of the information found in public registers and databases it is virtually impossible to co-ordinate this information. The Register of the Reporting Obligations of Enterprises provides precisely this overview, thus facilitating the co-ordination of public forms for business and industry.

The Register of the Reporting Obligations of Enterprises is operated by the Brønnøysund register centre.

7.1.2 SERES

SERES II is a project for the implementation of a national metadata register. SERES II is run by the Brønnøysund Register Centre. The project has an obligation to deliver solutions supporting the ongoing production of Altinn, especially the production for the Tax Directorate (SKD). Even if this requirement has constrained the project with respect to approach and priority, the project seeks to

develop concepts and a functional architecture that both satisfy the production of today's systems and the future requirements of coordination and cooperation within Civil Service.
The SERES solution shall in 2009 produce the following deliverables:

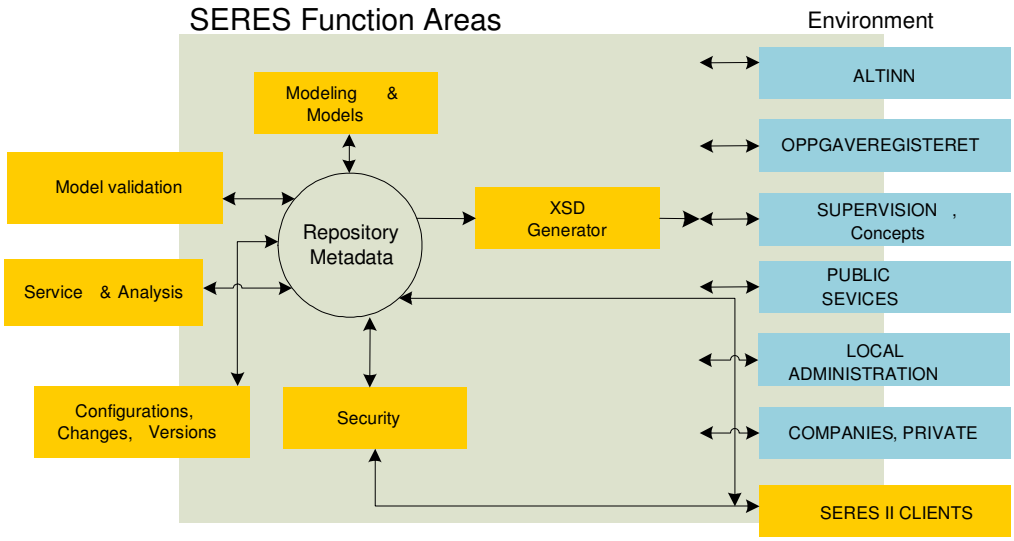
1. Reestablish necessary technical functionality such that The Register of the Reporting Obligations of Enterprises better can fulfil their assigned tasks.

2. Support the production of metadata for old and new Altinn, and the migration from the old to the new Altinn.

3. Support Altinn during the introduction of an increasing number of users.

4. By assessment of the current metadata, increase the quality of exiting data.

The SERES solution can by further modeling support new future tasks.
The following figure gives an overview of the functional areas of the SERES solution. The picture also indicates the cooperating environment of SERES. A good solution for storing and management of metadata (*Repository Metadata*) is the core of the solution. Various client tools, adapted to the users roles and requirements, use the data stored in the Repository Metadata. The other boxes in the figure offer various modeling functionality, functionality for quality assurance of models and data, functionality for searching, navigation, statistics and report generation, functionality for change management and version control and in addition solution for authorization and authentication of users.



7.1.3 Altinn

Altinn is a service for data collection to public sector from businesses and citizens. Nearly 30 public organisations use Altinn, e.g. Tax Authorities, the Brønnøysund Centre, Statistics Norway. Altinn is run by BR.

7.1.4 Volven.no

Source: http://www.volven.no/index.asp

Volven is a metadata service for health sector. It gives an overview and access to the common metadata for the health sector such as codes, classifications, terms, concepts, definitions etc. Volven is operated by KITH (Norwegian Centre for Informatics in Health and Social Care).

Volven (formerly also referred to as Balder) is established and maintained by KITH (Norwegian Centre for Informatics in Health and Social Care).

Volven (www.volven.no) is the metadatabase and the de facto publication place for metadata used in the Norwegian healthcare sector. The intention of Volven is to give a complete overview and access to up-to-date collections the unified metadata basis for the healthcare sector in Norway. The term "unified metadata basis" includes in this context coding schemes, classification systems, terms (terminology), concept definitions, data definitions etc.

The current implementation of Volven consists of, among other tings, the following types of metadata:

- Catalogues of concept definitions (in Norwegian: definisjonskataloger): Verbal definitions of concepts/terms that are commonly used in the Norwegian healthcare sector. One example of such definitions is: "A patient referral is a formal request to a healthcare institution for investigating or treating health problem(s) of a given patient" (freely translated from Norwegian).
- Administrative coding schemes: These are coding schemes for how to code various types of "administrative information". One example of such administrative coding schemes is: "Address type" and therein "H = Residence address; …; TMP = Temporary address; WP = Work address" (freely translated from Norwegian).
- Health professional coding schemes: Currently only a short description of each of the coding schemes that are used and the URL-link to where the whole coding scheme is published at KITH's web site.
- Messaging standards: KITH-standards for electronic messaging with and within the Norwegian healthcare sector. Examples of such messaging standards are: Application acknowledgement; Electronic prescription; Electronic referral; Electronic discharge letter; etc. Messaging standards established by KITH are documented by UML-based information/domain models in addition to XML-schema.

Functionalities that are yet to be implemented are, among other things:

- Querying and downloading all coding schemes that are used in a given message type (i.e., coupling between coding schemes and messaging standards mentioned above).
- Querying and downloading all messaging standards that use a given coding scheme (i.e., coupling between coding schemes and messaging standards mentioned above).
- Coding schemes that are published and/or maintained by other organisations than KITH, but are used within the healthcare sector (such as coding schemes maintained by NAV – the Norwegian labour and Welfare Administration, NPR – the Norwegian Patient Register, etc.).

Technically, Volven is implemented in conformance with ISO 11179 [ISO 2004]. Volven is equipped with a web service interface for integration with and access from external applications. The web service interface is e.g. in use with the solution at KITH for automatic testing and certifying vendor implementations of messaging standards when checking the validity of codes used in the test messages from the vendors.

Volven (at that time referred to as Balder) was one of the fem pilot projects of "eNorway 2002-2003 (electronic content)", which are referred to as best practice in publishing and managing metadata.

7.1.5 Norway.no (Norge.no)

Source: http://www.norway.no/omnorgeno/Default.asp?
Norway.no is the gateway to the public sector in Norway. The portal aims to help members of the public find public information and access public services more easily. Norway.no is a service run by the Agency for Public Management and eGovernment (DIFI).

7.1.6 MyPage (MinSide)

Source: http://www.norway.no/minside/
The idea with MyPage is to bring public service offerings together in a web portal where a citizen can get its own customized page. The information will be structured thematically, and the services will be grouped and sorted according to the citizens' needs. MyPage represents an opportunity for a simple and effective dialogue with the public sector. It will also give an overview of what information the individual agencies have on the citizen.

7.1.7 LOS

Source: http://www.norge.no/los/
LOS is a classification system for information about public services. The aim of LOS is to make it possible to share information about public services and to make it simpler for citizens to find their way in public sector.

7.2 Research

7.2.1 eGovMon - eGovernment Monitor

Source: www.egovmon.no
The eGovMon is an innovation project co-funded by the Research Council of Norway under the VERDIKT program. The project started in 2008 and will last for 3 years.

eGovMon will deliver a prototype implementation of a large scale, online observatory for benchmarking eGovernment services in four areas: accessibility, transparency, efficiency and impact. A set of well defined indicators will be identified for each area, using a coherent assessment methodology. To develop and maintain the measurement framework, an open, participatory and inclusive process will be established. Evaluation results will be gathered through automated tools, when possible, and supplemented by survey and manual assessments. All project results will be released under an open license, and all software will be open source.

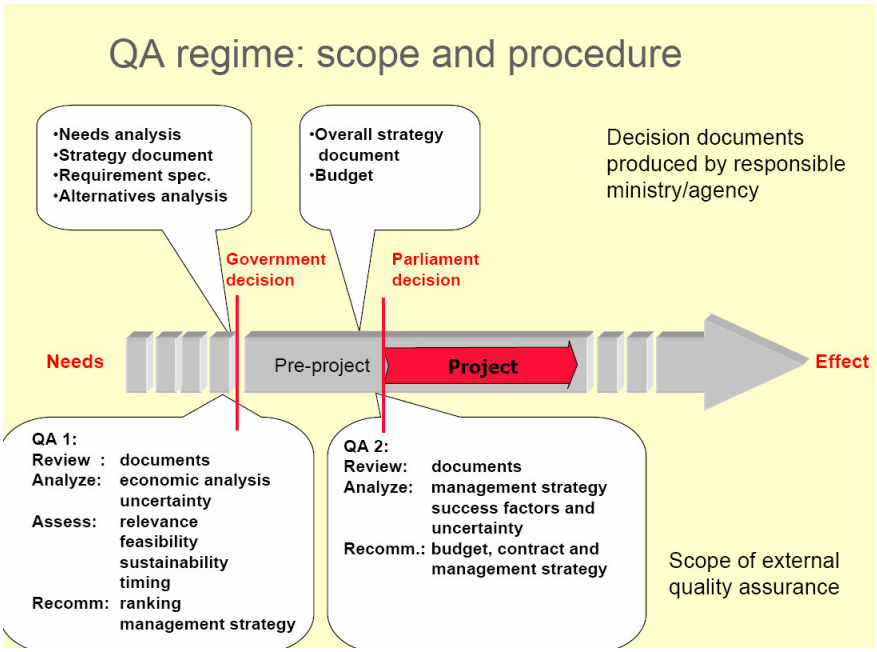
Relevance for Semicolon
One of the main goals of Semicolon is to develop metrics and indicators to measure the effects of interoperability. The eGovMon project may produce results of interest for the metrics-activity in Semicolon.

8 METRICS, ROI, OBSTACLES AND OTHER FUN STUFF

8.1 Quality Assurance of major Norwegian public investment projects

The purpose of the QA regime is to both assure proper choice of solution-concepts in large public projects and to assure that the project is under actual political steering. The chosen alternative shall, under given circumstances, use the available resources in a best possible way and create the best possible value-creation. The QA regime is used in project having an estimated cost of 500 million NOK or above

The illustration below shows the timeline and the two phases in the external QA of major Norwegian public investment projects. [Updates on project governance in Norway]



QA1: Selecting the right concept at a time when alternatives are available
QA2: Consolidating the choice of concept and corresponding budget

The ambition of the QA regime is as follows:
1. Increased awareness of QA in public sector
2. Improved procedures and practice in public sector
3. Trickle-down effects in private sector
4. More realistic budgets
5. Better budgetary compliance
6. Better choice of concepts
7. Higher benefit/cost ratio
8. More successful projects

In winter 2007/2008, the Altinn project was through a successful KS review.

This work is related to the Concept research programme. This programme focuses on Front-end Management of major investment projects. It aims to develop know-how to make more efficient use of resources and improve the effect of major public investments. It is financed by the Norwegian Ministry of Finance. [Concept research programme]

8.2 Interoperability barriers

Interoperability problems identified by Bekkers [Bekkers, 2007, p.377]:

- Administrative interoperability: Conflicting, exclusive or overlapping jurisdictions and accountability
- Legal interoperability: Different legal regimes with conflicting rights and obligations, e.g. in relation to privacy and safety regulations
- Operational interoperability: Different working processes and information processing processes, routines and procedures
- Technical interoperability: Incompatibility of specific 'legacy' ICT infrastructure (hard and software)
- Semantic interoperability: The idiosyncrasy of information specifications and the lack of common data definitions
- Cultural interoperability: Conflicting organisational norms and values, communication patterns, and grown practices

Interoperability barriers identified by Eynon-Margetts [Eynon-Margetts, 2007, p.5]:

- Leadership failures resulting in slow and patchy progress to eGovernment.
- Financial inhibitors limiting the flow of investment to eGovernment innovation.
- Digital divides and choices, where inequalities lead to differences in motivations and competences that constrain and fragment eGovernment take-up and fail to address particular user needs.
- Poor coordination across jurisdictional, administrative and geographic boundaries that holds back eGovernment networking benefits.
- Workplace and organisational inflexibility impairing adaptability to new networked forms of information sharing and service provision.
- Lack of trust heightening fears about inadequate security and privacy safeguards in electronic networks.
- Poor technical design leading to incompatibilities between ICT systems or difficult-to-use eGovernment services. Where such services lag behind innovative applications used by society more generally, government organisations will find it increasingly difficult to address issues of interest to online communities, which will tend to have different communication channels and mechanisms for producing content.

See also interoperability challenges identified by the Athena Interoperability Framework (Chapter 6.2)

8.3 Measurements and metrics

8.3.1 Requirements to measurement of interoperability

In order to measure anything, some kind of a model is required. The nature of such model(s) is however an open issue.

Semicolon has a focus on the ability to measure interoperability. This implies that

- we need to be able to measure the degree of interoperation, and perhaps the quality, characteristic and/or utility (individual, mutual and collective) of the interoperation
- we need to be able to measure the ability to establish interoperation in whatever situation might arise in the future, - "quickly and cheap". This implies an ability to change.

If we want to lay the ground for measuring the development of the interoperability in the Norwegian Public Sector over years, the model we employ must also endure and adapt to changes in the real world. Ideally, we should measure on models that exist as a result of actual work, - not models that are created for the sake of measurement, - or models expected to be maintained by people outside of the core processes. One modelling team cannot be expected to successfully reflect the business of the 700 governmental bodies, rather should one look for ways of utilizing models that are produced by the bodies themselves. Since there are currently neither common models nor common operational platforms, the question of how to achieve this is a core initial question. Given a common platform, it is possible to define algorithms that analyse metadata in order to give some measure of interoperability, - either individually, bilaterally (between two potential partners) or collectively as a whole. It would then also be possible to impose new questions in hindsight upon the same models, such as evaluating interoperability on partial domains or in given contexts.

Is the establishment of common platforms where the bodies can maintain their characteristics a realistic path, can we utilize the information published by the 700 bodies as is, - or will we eventually have to fall back on some central team (Semicolon?) that completes both modelling, measurement and evaluation? This is an important area that needs to be addressed in Semicolon.

8.3.2 Requirements to approach for interoperability

The federated history of Information Technology has led the thinking within the field of interoperability to motivate developers of systems to make systems that are built in common ways so that they become interoperable "by design". ISO14258 (Requirements for Standards and model interoperability) distinguish between the following three forms:

- Integrated – use of common modelling form
- Unified – use of common meta-level structure
- Federated – use of model characterization to dynamically accommodate different models.

None of the above does however seem to characterise the star of most recent development: The operational platform itself. Google has now for a couple of years provided an operational platform for sharing mail, calendars and online documents with advanced change management features. Facebook offers whoever wants to connect to build upon the core FOAF (Friend of a friend) model and write their own applications within the same operational platform. Freebase offers a general modelling platform, - available for anybody to make their contribution in context of other peoples work both with regards to metamodels, metadata and real-world information reflecting OMGs M0 metalevel, - with an open API that allows anyone to build their own user interface or analyzing algorithm specifically for a given view, context or purpose. These are not utilizing architectural frameworks to ensure interoperability (like MDD/UML, TOGAF, Zachman, FEAF, OIO or others) through some common design tool or common language, - they simply "do it" by providing an operation central platform where the actors are encouraged to participate, - and where the operational needs and cost of scalability, backup, redundancy, 7/24 operation, etc. is provided by the operator. Operational interoperation is immediately available and testable, - and the need to build support for quality assessment, dependability, promotion, partial utilization, and quality improvement is a research theme.

Wikipedia and Wikis are other "Web2.0" platforms that promise quicker and cheaper interoperability by business people and non-IT actors, - without the need for the support of any functionality to pass through some IT software development cycle.

8.3.3 Cost-benefit analysis / methods / metrics

8.3.3.1 Norwegian government guidelines

Back in 1977 the Norwegian Ministry of Finance published the first guide to economical analysis of the Norwegian society. This guide has since been revised a number of times and the last version was published in 2005.

January 1st 2004 the Norwegian Government Agency for Financial Management (SSØ) was established by the Ministry of Finance. This agency is now responsible of administering the economical guideline and serving governmental agencies in how to understand and do their cost – benefit analysis based on the guideline.

The guide describes three different types of cost–benefit analysis:

1. **Benefit-cost analysis:** A systematic survey on advantages and disadvantages given a specific initiative. The benefits and costs are given in Norwegian crowners (NOK) as far as possible.
2. **Cost efficiency analysis**: A systematic evaluation of costs given different alternative initiatives to obtain a given objective. The cost is then given in Norwegian crowners (NOK) and one will strive to find the most affordable alternative to obtain the given objective.
3. **Cost – impact analysis**: A survey on costs of different initiatives that address the same problem, but the impact may differ between the different initiatives. The different impacts must then also be considered before one may conclude on the best initiative.

The Norwegian Government Agency for Financial Management has developed a more explicit model based on the guideline from the Ministry of Finance and is offering this model to other governmental agencies together with support on how to make use of this model.

8.3.3.1.1 Challenges

There are at least two issues to be addressed when discussing possible improvements to the guideline from the Ministry of Finance or the model from the government agency for Financial Management:

Cross governmental initiatives: How to analyse initiatives where the costs and benefits do not occur within the same jurisdictional area? How to make the decision rationale?

Benefits that cannot be quantified: The way of evaluating the value of a quality that cannot be measured in either time or money.

The Semicolon project addresses several such initiatives/cases that will challenge the way we consider the cost – benefit aspects. Several of the governmental partners in Semicolon are expecting Semicolon to contribute to handling these issues in a more predictable manner.

9 LAW AND LEGAL CONDITIONS

Norwegian public sector is controlled by legal regulations given by Norwegian Law. The transition from regular government to e-Government has to be in compliance with the law and is therefore being evaluated with respect to all legal aspects such as data privacy and security requirements.

Traditionally the legal department has been involved in considering possible implications of developing new legacy systems or any ICT system supporting the public sector business. Professor Dag Wiese Schartum at The University of Oslo, Faculty of Law, Section for Information Technology and Administrative Systems has considered the fact that there are no ICT systems supporting the law and the development of the law itself. This fact explains how some descriptions/definitions have been defined in different ways in the law, even though the meaning and comprehension generally would be the same.

10 BEST PRACTICE INTERNATIONAL

10.1 UN eGovernment survey 2008

The UN eGovernment survey 2008 [UN eGov survey, 2008] has the subtitle “From E-Government to Connected Governance”. It focuses on the evolvement of E-Government; from traditional government via E-Government to “Connected government”.

One of the measurements done in the 2008 survey is based on an index called the e-government readiness index. This is a composite index comprising the web measure index (see figure xx below), the telecommunication infrastructure index and the human capital index.

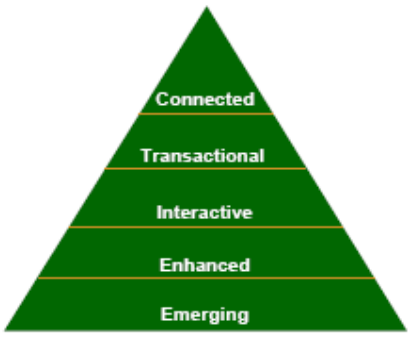


Figure: Phases of the Web measure index

10.1.1 E-Government readiness rankings

The world average of the global e-government index continues to increase as more countries invest resources in developing websites that are informative. Most countries have e-information on policies, laws and an archive section on their portals/websites.
The gap between e-information, e-consultation and e-decision-making is still wide for developing and developed countries.
For the first time since this survey has been produced, there is a new leader. In the 2008 Survey, **Sweden** (0.9157) took the number one spot from the **United States**. The Scandinavian countries took the top three spots in the 2008 Survey, with **Denmark** (0.9134) and **Norway** (0.8921) in second and third place respectively. The **United States** (0.8644) came in fourth.

TECHNICAL REPORT

In this year's e-government readiness rankings, the European countries make up 70 per cent of the top 35 countries. The Asian countries make up 20 per cent of the top 35 and the North American and Oceania regions 5 per cent. The European countries as a group have invested heavily in deploying broadband infrastructure, coupled with an increase in the implementation of e-government applications for their citizens. Yet, according to the ITU, the European countries make up nine of the top ten countries in broadband subscribers per hundred, with **Denmark**, the **Netherlands** and **Iceland** being the top three countries.

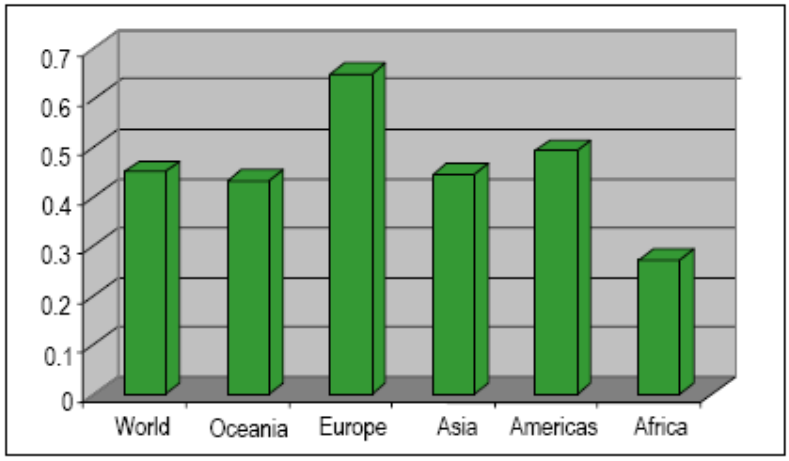


Figure: Regional Average of e-Government Readiness

Figure xx clearly shows the difference between the five regions, with Europe (0.6490) having a clear lead over the other regions, followed by the Americas (0.4936), Asia (0.4470), Oceania (0.4338) and Africa (0.2739). Asia and Oceania are slightly below the world average (0.4514), while Africa lags far behind.

Rank	Country	E-Government Readiness Index
1	Sweden	0.9157
2	Denmark	0.9134
3	Norway	0.8921
4	United States	0.8644
5	Netherlands	0.8631
6	Republic of Korea	0.8317
7	Canada	0.8172
8	Australia	0.8108
9	France	0.8038
10	United Kingdom	0.7872
11	Japan	0.7703
12	Switzerland	0.7626
13	Estonia	0.7600
14	Luxembourg	0.7512
15	Finland	0.7488
16	Austria	0.7428
17	Israel	0.7393

Rank	Country	E-Government Readiness Index
18	New Zealand	0.7392
19	Ireland	0.7296
20	Spain	0.7228
21	Iceland	0.7176
22	Germany	0.7136
23	Singapore	0.7009
24	Belgium	0.6779
25	Czech Republic	0.6696
26	Slovenia	0.6681
27	Italy	0.6680
28	Lithuania	0.6617
29	Malta	0.6582
30	Hungary	0.6485
31	Portugal	0.6479
32	United Arab Emirates	0.6301
33	Poland	0.6117
34	Malaysia	0.6063
35	Cyprus	0.6019

Table: Top 35 Countries in the 2008 e-Government Readiness Index

It is worth noting that in the 2008 Survey, there are no countries in the top 35 from the African, Caribbean, Central American, Central Asian, South American and Southern Asian regions. The high cost of deploying a robust infrastructure capable of handling egovernment applications is one reason for this discrepancy. In addition, many developing countries have been unable to fully implement their e-government policies, mainly due to other competing pressing social issues that need to be dealt with in the context of tight budget constraints, such as: health, education and employment, to name a few.

Northern Europe

The Northern European region was the strongest region in Europe. Sweden (0.9157), Denmark (0.9134), and Norway (0.8921) were the top three countries in the Survey and were standouts on the web measure in 2007-2008, with Denmark and Sweden ranking number one and two, and Norway finishing fourth. Sweden with its newly revamped eservices portal http://www.sverige.se, 'your guide to Sweden's public sector', and Norway with its redesigned primary site http://www.regjeringen.no have improved from previous years, but Denmark http://borger.dk/ still leads the way among the Scandinavian countries and globally.

Interestingly, the Scandinavian countries surveyed all employ a similar web strategy. They each have a primary site that is informational and a tightly integrated, gateway site for e-services. Using this approach, each of the Scandinavian countries scored very high on the availability of services and transactions, the clear area where they excelled compared to most other countries. Compared to previous years, Denmark, Norway and Sweden still have a large amount of content available in other languages, but not nearly as large a percentage as in the past. This is mainly due to the enormous growth of the information and content available on their sites.

Another feature the Scandinavian sites have in common is that they include quite a bit of information about and references to, e-participation. All established e-participation or edemocracy commissions and they have all sorts of publications and findings.

Country	2008 Index	2005 Index	2008 Ranking	2005 Ranking
Sweden	0.9157	0.8983	1	3
Denmark	0.9134	0.9058	2	2
Norway	0.8921	0.8228	3	10
United Kingdom	0.7872	0.8777	10	4
Estonia	0.7600	0.7347	13	19
Finland	0.7488	0.8231	15	9
Ireland	0.7296	0.7251	19	20
Iceland	0.7176	0.7794	21	15
Lithuania	0.6617	0.5786	28	40
Latvia	0.5944	0.6050	36	32
Region	0.7721	0.7751		
World	0.4514	0.4267		

Table: E-Government Readiness for Northern Europe

Over the past year, the United Kingdom has revamped its government online system, through an initiative to pare down the numerous (hundreds) of government websites available to the public. The UK's main government portal, www.directgov.uk, was redesigned in 2008, which appears to have resulted in a drop in the web measure rankings for the UK. Yet, the Directgov.uk website ('Public service all in one place') does one of the best jobs of joining up information and services from the central government as well as local authorities. The main site is filled with information, and has a consolidated directory and services listing for the central government with local authorities, along with additional excellent linkages to local government services and resources. Furthermore, the main site has new citizen communications features, such as a mobile government portal, as well as a separate business gateway www.businesslink.gov.uk. While the UK national site may have slipped in the rankings, the site with its comprehensive information and services covering different levels of government still provides good value to the citizen user.

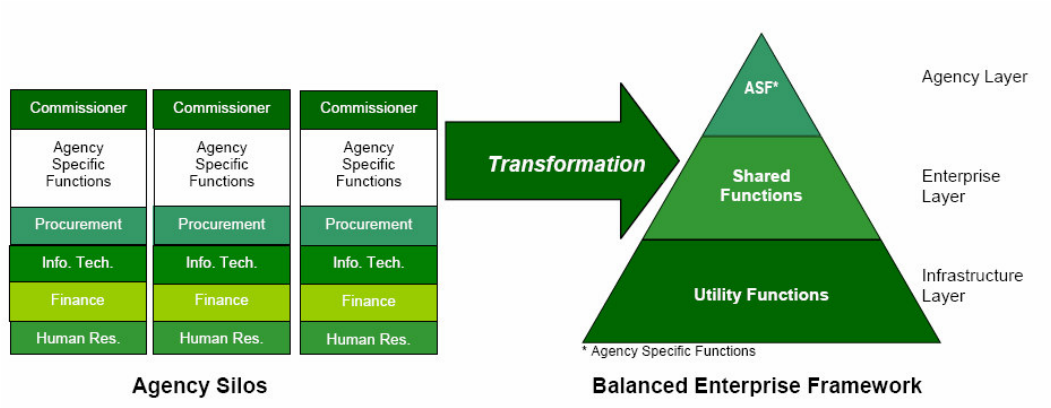
It should be noted that all the countries in Northern Europe are in the top 20 per cent of the infrastructure, education and web measurement indices, as well as the egovernment readiness index.

The Social Security and Labour Ministry of Lithuania http://www.socmin.lt/ provides an online English-Lithuanian dictionary of social terms, employment and health topics. It also gives advice on how to find work, with the option of consulting with a government official via the Internet.

10.1.2 PartII: "From e-Government to Connected Governance"

Part II of the report addresses the issues of moving from the siloes of eGovernment to a more holistic service of "connected Governance", where all back office systems are or can be integrated.

The following figure (Source: State of Minnesota, 2005) illustrates the point:



The refer to the 2007 Accenture global review, which states:

"After years of focusing primarily on the front-end (highly visible, citizen-facing aspects of service delivery), governments are now trying to take a more holistic approach. While they are still trying to bring things together for citizens at the front office, they have come to the point where they also need concrete plans for making a superior front-end customer experience operational on the backend. In short, this means a renewed emphasis on the infrastructures and workforce that will be able to take the promise of citizen-centred service through to practice."

The report discusses

- multiple conceptual organizational models for integration (they talk of collaboration and integration, not so much "interoperability"),
- "wikinomics – how Mass Collaboration Changes Everything", web2.0 and eDemocracy, building on the new possibilities and forms of collaboration on the web, - and retting it up against the traditional and strong governmental needs for security and control
- Drivers for integration and "delivery" in horizontal/vertical connection points, operational and strategic integration levels.

It also discussed Centralized versus Distributed control, where the following paragraph proposes a direction:

> "The notion of control is fundamental here to understanding the reframing that must occur. All organizations and institutions require some form of control, but the widening interest in new governance systems is testament to the need to view control as less a means to shape every aspect of behaviour (i.e. process control) and more a basis for coordinated and shared actions orchestrated on the basis of outcomes and objectives"

Its discussion of "Origins and Emergence of e-Government" concludes on direction of inclusion without giving the answers to "how":

> E-government as a transformational project should be framed first and foremost as a conversation, one that should ideally resonate across the widest possible set of individuals and organizational actors within any given jurisdiction. The likeminded global challenge for the world as a whole is to extend this conversation to a transnational plain in a manner that enables a greater exchange of knowledge and resources globally, and a more informed and well-devised set of e-government strategies nationally and locally.

Relevance for Semicolon

This part of the report is highly relevant for semicolon, by raising relevant issues, drivers and insights/opinions that go beyond or across the more topic-oriented state of the art that a document like this could be oriented towards.

They identify "two fundamental questions", that are in line with the work that we undertake in Semicolon, and show that we are "on the right path":

> From the perspective of more horizontal but in reality networked governance solutions that are the essence of service transformation and effective security strategies, the two fundamental questions that remain stubbornly unanswered include:
>
> - How to motivate public managers to share data and, more generally, to work jointly for the public good; and
> - How to understand and influence the range of barriers, from psychological and social to structural, political and technical, that mitigate across cross-agency initiatives

These questions are to be addressed in Semicolon.

10.2 Denmark

Source: http://modernisering.dk/da/projektside/
Denmark has had an eGovernment project since 2001. For the period 2007 – 2010 they have defined a strategy with three elements:

1. Better digital service
   Development of user-friendly digital services in areas with largest impact for citizens, businesses and public sector itself
2. Increased efficiency in public sector
   Reallocate resources from administration to user services
3. Increased collaboration
   Binding collaboration across organisations to reduce double work, increase reuse and ensure interoperable systems

Relevance for Semicolon
Danish strategies are in the midst of Semicolon topics.

10.3 Australia

Source: http://www.sbr.gov.au/content/default.htm

Standard Business Reporting is an initiative from the Australian Government. It aims to reduce the reporting burden for businesses.

Relevance for Semicolon
Methods and results should be evaluated by Semicolon.

10.4 Research

The European Unions 6th Framework Programme (FP6) for Research had considerable focus on Enterprise Interoperability (EI), motivated partly from and building on the roadmap delivered by the the FP5-project IDEAS (http://interop-vlab.eu/ei_public_deliverables/ideas-deliverables/).

Most notably, and with relevance for Semicolon, - we find the projects Athena, Genesis, Fusion and Abilities, of which we include summaries over the three former below

Currently, the EU "Framework Programme 7 for Research and Technological Development" (http://cordis.europa.eu/fp7/) is under execution, partly motivated by the EI roadmap [Cordis EI Roadmap, 2006] delivered by the Athena Project in FP6.

We note the "Grand Challenges" identified by this roadmap:

Grand Challenge(1): Interoperability Service Utility

"This Roadmap envisions a diversity of continuously evolving ecosystems of enterprises. Interoperability of enterprises will be a key feature both within and across such ecosystems. Specifically, interoperability will be a **utility-like** capability for enterprises, a capability that is:

- Available at (very) low cost
- Accessible in principle by all enterprises (universal or near-universal access)
- "Guaranteed" to a certain extent and at a certain level in accordance with a set of common rules
- Not controlled or owned by any single private entity."

Grand Challenge(2): Web Technologies for Enterprise Interoperability

The roadmap identifies the web as the future platform ("operating system") of future Enterprise interoperability, where actors will need to adapt or die - but argues that as such it is immature and needs improved "Operating System" functionality.

Grand Challenge(3): Knowledge-Oriented Collaboration

Sharing Best Practices and experiences of collaboration, for further elaboration, - please see the reference above.

Grand Challenge(4): A Science Base for Enterprise Interoperability

To illustrate, we site an interesting and point-making paragraph from the Problem Statement of this challenge: "The market is saturated with technology-based solutions that claim to support interoperability for
enterprises. It also shows a profusion of interface standards and specifications that make the same claim. Yet, enterprises still cannot exchange information easily and transparently. There are several reasons for this.
Traditionally, Enterprise Interoperability solutions are linked closely to specific market sectors, application areas, and technology trends. They generally work well within the particular, self-defined, static environment for which they were designed. However, they cannot be modified easily to deal with changing technologies. Moreover, the solutions from one provider are, frequently, incompatible with

the solutions from another. More specifically, they are designed to meet developers' needs rather than users' needs (see Grand Challenge in Chapter 5), they are incompatible with the emerging Web technologies (see Grand Challenge in Chapter 6), and they do not support enterprise-wide decision making (see Grand Challenge in Chapter 7).

On the research side, the central problem is that researchers, even if they are one step ahead of the providers, are also technology bound. The existing concepts, methods and techniques are inadequate because they resolve the same problems using the next generation technology. This approach cannot keep pace with the demands of enterprises as they collaborate and compete within a process of Enterprise Interoperability research roadmap 31 July 2006 Page 35 / 45
technology-enabled *business* innovation that is transforming the rules of the game and the nature of business. With a view to the future, Enterprise Interoperability research needs to break away from specific technology boundaries and be based on solid and rigorous scientific theories and principles.

Although it is too early to learn from these projects' deliverables, their focus and courses of action are of relevance to Semicolon. Following the FP6 projects, we therefore include a section on current relevant projects of FP7.

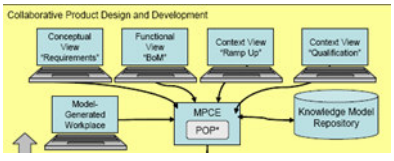
10.4.1 ATHENA[3] (EU FP6)

Source: http://interop-vlab.eu/ei_public_deliverables/athena-deliverables

We give the Athena project considerable space in this report, - because we see the project, its broad, multidisciplinary approach and contributions, to carry significant impact to the research on interoperability within the recent 5 years.

The Athena project (2003-2007) addressed technologies for interoperability by merging three different disciplines (Architecture and Platforms, Enterprise Modeling and Ontologies) in an intended holistic approach. In particular, the project addressed the following areas, where an important common characteristic is that they are model-driven:

- A1: Enterprise Modeling in the context of Collaborative Enterprises – addressed the challenge resulting when the collaborating enterprises choose different tools and languages to describe themselves and their common collaboration models. A flexible common language POP* and a collaborative modeling platform MPCE[4] was developed or elaborated. MPCE supports MUPS (Model-Configurated, User-composed Platforms and Services) in an integrated operational platform for both definition/communication of work and operation/"execution" of the definitions (Smith-Meyer 2005) on an operational platform for work and process support. The type of collaborative modeling in Athena comes from the case where a product is developed collaboratively, and the value of potential transfer to Governmental distributed modeling is an issue to be explored. Iterative and partial model import/Export functionality (including identity management across units) and the need to highlight added and deleted elements were given solutions by this

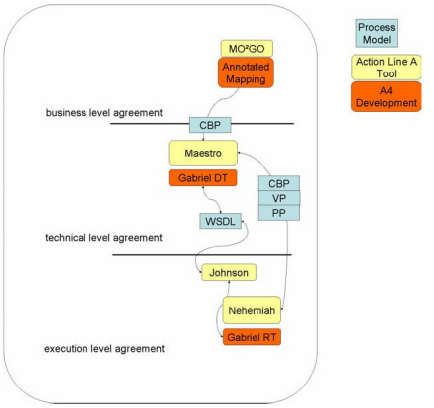


[3] Advanced Technologies for Interoperability of Heterogeneous Enterprise Networks and their Application

[4] MPCE – Modeling Platform for Collaborative Enterprises [Athena DA4.2]

Athena Project.

- A2: CBP – Cross-Organizational Business Processes – provided and extended both a modeling platform and parts of a execution platform that makes a distinction between ”private processes” and public processes, where the private processes hides the details of how the collaborating partners performs their tasks and responsibilities in a common public, shared process. The figure shows the relevant Athena tools at the level of business agreement, technical agreement and execution agreement



- A3: Knowledge Support and Semantic Mediation Solutions – uses ontologies to resolve semantic challenges between the interoperating partners, based on ATHOS (Athena Ontology Management System). Satellite languages are resolved towards a central ontology. OPAL (Object Process Actor Language) semantically describes the core elements of the ontology, showing a work-oriented context backbone.
- A5: “Perspectives in Service-OrientedArchitectures and their Application inEnvironments that Require Solutions to be Planned and Customizable” – built on the challenges and possibilities of a Peer-2-Peer platform
- A6: Model-Driven and adaptive interoperabilities – built on OMGs Model-Driven Architecture (MDA) and Development (MDD) and their concepts of Conceptual Models, Platform Independent Models (PIM) and Platform Specific Models (PSM) to provide a number of transformations tools, methodologies and guidelines, a Platform Independent Model for Service-Oriented Architectures (PIM4SOA) and an Execution Framework, - all in relation to the (open) Eclipse software platform. The concept is dominated by the goal of producing software, but is not limited to this scope at conceptual level, even if it is hard to separate concerns and not to fall into “the software trap”.

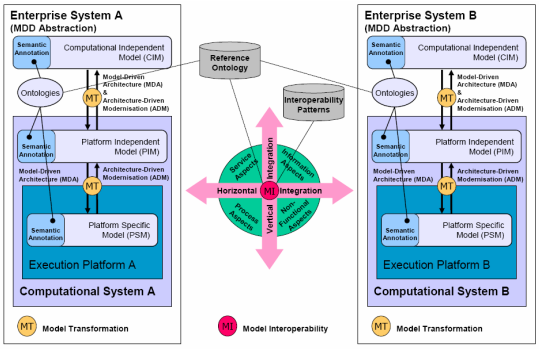
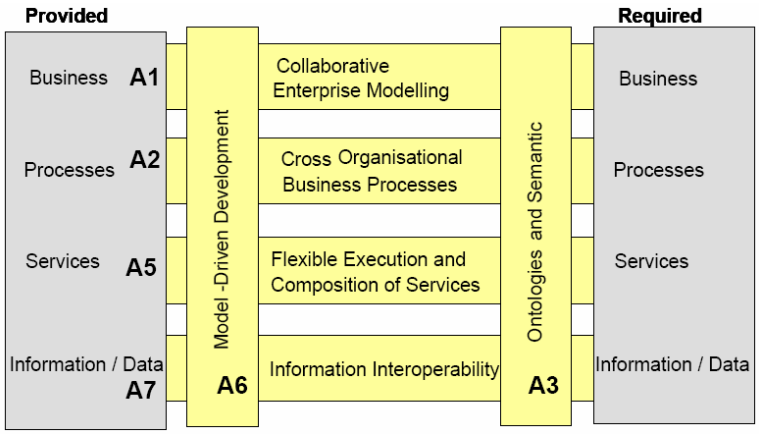


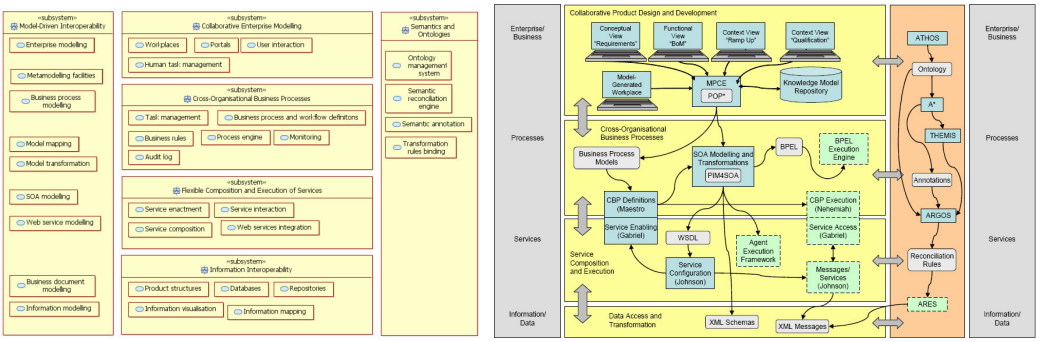
Figure 121: Reference model for conceptual integration

- Then, A4 brings the above together, as basis for the Athena Interoperability Framework

With reference to the Athena Interoperability Reference Architecture, their contributions were placed holistically as shown in the following figure:



The Main functionality offered, together with example tool realization and their interconnection is shown in the following figure.



"Main functionality offered by the ATHENA platform" and an "Example of possible configuration and implementation of the ATHENA platform" (from Athena DA4.2, figure 54)

It is worth noting, however, that even if some of these tools are built upon commercial platforms, the actual solutions are still research products without the characteristics one expect from commercal tools and platforms. Contributed Solutions can be found at http://modelbased.net/aif/solutions/index.html

10.4.2 GENESIS (FP6)

Source: http://interop-vlab.eu/ei_public_deliverables/genesis-deliverables
GENESIS (2006-2008 EU-funded FP6-IST- 027867) addresses "Enterprise Application Interoperability via Internet-Integration for SMEs, Governmental Organisations and Intermediaries in the New European Union. The main target of the GENESIS project is the research, development and initial, precompetitive application of the needed methodologies, infrastructure and middleware software components that will allow the typical, usually small and medium, European enterprises to conduct Business transactions over Internet, by interconnecting its main transactional software applications and systems with those of collaborating enterprises, banking/social insurance institutions and governmental bodies, with respect to the evolving legal and regulatory status"

Relevance to Semicolon:

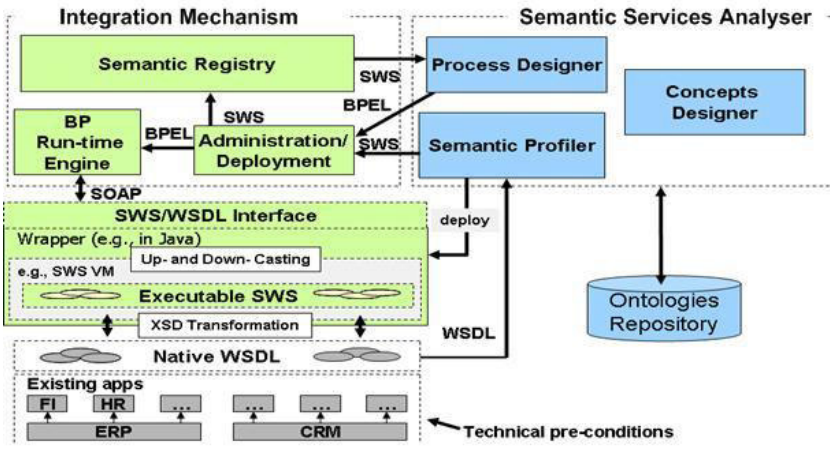
The Genesis project shows high relevance for Semicolon in the justice domain, justifying further study. It studies the income tax laws and in a number of EU states, thereby also giving input to a "side-case" of semicolon through the eDialog project.

10.4.3 FUSION (FP6)

Source: http://interop-vlab.eu/ei_public_deliverables/fusion-deliverables

FUSION (2006-2008 EU-funded FP6-IST- 027385) "aims to promote efficient business collaboration within enterprises (incl. SMEs) by developing technologies for the semantic fusion of heterogeneous businesses applications. Intercultural and regulatory aspects of the enlarged Europe countries are considered instrumental in the FUSION solution. It utilizes BPEL4WS SA-WSDL and Semantically-Enriched UDDI and OWL technologies to build

- a Semantic Registry that facilitates categorization, search and discovery services that enable the semantic discovery of published SA-WSDL compliant Service Profiles from design-time components
- A Business Process Execution Environment (ActiveBPEL) on a native BPEL4WS engine that takes care of persistence, queues, alarms and other execution details.
- An administration component that serves as a mediator among a Semantic Services Analyzer and the integration mechanism



(The Global FUSION Reference Architecture – from http://interop-vlab.eu/ei_public_deliverables/fusion-deliverables/wp3/, page 15)

10.4.4 Current EU Research focus in Research Framework Programme 7 (FP7, 2007-2013)

The EU's 7th Framework Programme for Research and Technological Development (FP7) lasts from 2007 to 2013.

The FP7 projects are parallel in time too Semicolon and are more relevant to Semicolon through their focus, intentions and directions than through deliverables. As the name of FP7 suggests, we might expect a focus weighted more on the side of technological development than in Semicolon, but this may also be seen as development of the technological platform and operational infrastructure needed to support other or related research foci.

Under the ICT theme/programme (http://cordis.europa.eu/fp7/ict/ and http://ec.europa.eu/information_society/tl/research/documents/fp7-ict-4poverview.pdf) we find the projects of most interest to Semicolon, some of which are identified and described below.

Before we go there, however, - it is worth noting Semicolon-relevant goals, such as

- "making Europe's large public sector more efficient, and modernising sectors ranging from education to energy;"
- "tackling social challenges, improving quality of life and meeting the challenge of an ageing society."

- …and two of the seven areas of focus:

1: Laying Tomorrow's Networks

Tomorrow's information infrastructure will connect together billions of people, countless organisations and literally trillions of devices - PCs, mobile phones, servers, sensors and much more. This infrastructure will underpin economic development in all EU regions, and will be at the origin of new services and business opportunities throughout the economy.

Mastering the development of this infrastructure is essential to reaping the benefits of ICTs in areas as diverse as manufacturing and home healthcare. It is an immense challenge to make this network and service infrastructure more robust, resilient and secure.

5: A Healthcare Revolution

Sustaining Europe's healthcare systems is a major challenge, with healthcare already accounting for around 9% of EU GDP. The health sector is information intensive, so e-Health is emerging as an important new industry, with e-Health spending predicted to account for around 5% of the total health budget by 2010.

Research under this Challenge will improve the quality, availability and effectiveness of healthcare by developing ICTs to improve everything from healthcare administration to biomedical imaging, from personalised, home-based care to the creation of new medicines.

We find the following FP7-projects relevant, and have included parts of how they describe the situation (state of the art or future need) and mission/goal/direction of work:

10.4.4.1 COIN

Title: Collaboration and interoperability for networked enterprises
Research area: ICT-2007.1.3 ICT in support of the networked enterprise
Project start date: [2008-01-01]
Semicolon Relevance: Federated organization, adaptation as means for interoperability, agility, Model based

Description
"By 2020 enterprise collaboration and interoperability services will become an invisible, pervasive and self-adaptive knowledge and business utility at disposal of the European networked enterprises from any industrial sector and domain in order to rapidly set-up, efficiently manage and effectively operate different forms of business collaborations, from the most traditional supply chains to the most advanced and dynamic business ecosystems.

The mission of the COIN IP is to study, design, develop and prototype an open, self-adaptive, generic ICT integrated solution to support the above 2020 vision, starting from notable existing research results in the field of Enterprise Interoperability (made available by the Enterprise Interoperability DG INFSO D4 Cluster and specifically by the projects ATHENA, INTEROP, ABILITIES, SATINE, TRUSTCOM) and Enterprise Collaboration (made available by projects ECOLEAD, DBE, E4 and ECOSPACE)."

10.4.4.2 CUMMIUS

Title: Community-based ***Interoperability*** utility for SMEs
Research area: ICT-2007.1.3 ICT in support of the networked enterprise
Funded under 7th FWP (Seventh Framework Programme)
Project start date: [2008-02-01]
Semicolon Relevance: SME focus with relevant challenges to those of federated organisations, zero-cost of interoperation entry breaks barriers and supports opportunities

Description
"More than 99% of European enterprises are SMEs. While collaboration with other enterprises provides potential for improving business performance, enterprise interoperability research is yet to produce results which can be used by SMEs without the need for high start-up costs (learning costs as well as system purchasing and installation costs).

For a solution to be taken up by SMEs it must offer both 'zero costs of entry' and 'zero time for set-up' to the SMEs; its initial interaction with the system should follow familiar interaction patterns based on existing tools such as email or a web browser. Commius aims to deliver an adaptable and customisable software prototype, providing SMEs with 'zero-cost of entry' into interoperability using the ideas behind the Interoperability Service Utility. This will be made possible by a number of innovative scientific, technical and business advances over the existing state-of-art."

10.4.4.3 DIVA

Title: Dynamic variability in complex, adaptive systems
Research area: ICT-2007.1.2 Service and Software Architectures, Infrastructures and Engineering
Project start date: [2008-02-01]
Semicolon Relevance: Federated organization, adaptation as means for interoperability, agility, Model based

Description
"Context aware software systems that can automatically adapt to changes in their environments play increasingly vital roles in society's infrastructures. Such systems are called adaptive systems. The demand for adaptive systems appears in many application domains ranging from crisis management applications such as disaster management, first-response, transportation control, and power management to entertainment and business applications such as mobile interactive gaming, tourist guiding and business collaborations (e.g., through virtual organizations and dynamic service compositions). This demand is accentuated by the mobile and nomadic nature of many of these domains. The IDC analysts forecast a global increase in number of mobile workers to the level higher than 850 million by 2009.

The goal of this project is to provide a tool-supported methodology and frameworks for managing dynamic

variability of co-existing, co-dependent configurations in adaptive systems. This will be addressed through a combination of aspect-oriented and model-driven techniques."

10.4.4.4 MASTER

Title: Managing assurance, security and trust for services
Research area: ICT-2007.1.4 Secure, dependable and trusted infrastructures
Project start date: [2008-02-01]
Semicolon Relevance: Federated organization, need for trust, and measures

Description
"The business of the future will be characterized by highly dynamic service-oriented architectures where outsourcing and distributed management constitute the norm rather than the exception with an increasing complexity in security and trust requirements from regulations and business standards. Best-effort security will no longer be accepted and business entities will have to provide certified assurance services to customers and expect assured services from contractors in order to manage the associated business and technology risk.

MASTER aims at providing methodologies and infrastructures that facilitate the monitoring, enforcement, and audit of quantifiable indicators on the security of a business process, and that provide manageable assurance of the security levels, trust levels and regulatory compliance of highly dynamic service- oriented architecture in centralized, distributed (multi-domain), and outsourcing contexts.

To this extents MASTER will identify new innovation components in terms of key assurance indicators, key security indicators, protection and regulatory models and security model transformations coupled with the methodological and verification tools for the analysis and assessment of business processes. It will further define an overall infrastructure for the monitoring, enforcement, reaction, diagnosis and assessment of these indicators centralized, distributed (multi-domain), and outsourcing contexts. It will show a proof-of-concept implementation in the challenging realms of Banking/Insurance and in the e-Health IT systems.

MASTER will thus deliver a strategic component of the security and trust pillar of the European Technology Platform NESSI which makes it a NESSI strategic project."

10.4.4.5 NESSI 2010

Title: NESSI 2010
Research area: ICT-2007.1.2 Service and Software Architectures, Infrastructures and Engineering
Project start date: [2008-04-01]
Semicolon Relevance: Infrastructure for interoperability, - perhaps a bit on the technical side wrt Semicolon?

Description
"NESSI 2010, a SA funded in the context of FP7, focuses on specific tasks in support of the NESSI ETP, the Networked European Software and Services Initiative. NESSI was launched by industry in September 2005. Two years on, NESSI unites a community of 22 partners and 300 members around a single strategic objective: supporting the evolution from software to services. NESSI has delivered its Strategic Research Agenda and defined its delivery as NEXOF, the NESSI Open Service Framework.

NESSI intends to deliver a reference implementation of NEXOF and at the same time foster multiple implementations in open source or commercial models that deliver the required quality level of security, identification, interoperability and compatibility. NEXOF is key in ensuring that business services can be developed and made accessible to the largest possible audiences while relying on an underlying framework that is neither proprietary nor company centric.

In order to sustain the success of NESSI and validate NEXOF, it is time to enter the next stage where NESSI's Community efforts focus on increasing memberships in selected directions, supporting the research coordination role devoted to active ETPs and sustaining and increasing the dynamic momentum.

In this context, NESSI 2010 has been allocated specific tasks, focusing on strategy formulation, community clustering, coordination with national programmes, fostering involvement of ICT SMEs and

increasing awareness of NESSI through events and dissemination.

NESSI 2010's implementation is shared between a subset of NESSI partners that have committed to allocate the human and financial resources to the benefit of the entire NESSI community. These partners with their assigned primary responsibilities are Atos Origin (Interlinking activities), Engineering (NESSI Strategy & Communication activities), Thales (NESSI Office / Organisation strategy) and TIE (ICT SME specific activities). "

NESSI Web site: http://www.nessi-europe.eu/

10.4.4.6 OASIS

Title: Open architecture for accessible services integration and standardisation
Research area: ICT-2007.7.1 ICT and ageing
Project start date: [2008-01-01]
Semicolon Relevance: Connecting services in a user-friendly environment (for the elderly!) just like the Norwegian eDialog-concept, - but also with focus on services integration and seamless connectivity.

Description
"OASIS introduces an innovative, Ontology-driven, Open Reference Architecture and Platform, which will enable and facilitate interoperability, seamless connectivity and sharing of content between different services and ontologies in all application domains relevant to applications for the elderly and beyond. The OASIS platform is open, modular, holistic, easy to use and standards abiding. It includes a set of novel tools for content/services connection and management, for user interfaces creation and adaptation and for service personalization and integration.

Through this new Architecture, over 12 different types of services are connected with the OASIS Platform for the benefit of the elderly, covering user needs and wants in terms of Independent Living Applications (nutritional advisor, activity coach, brain and skills trainers, social communities platform, health monitoring and environmental control), Autonomous Mobility and Smart Workplaces Applications (elderly-friendly transport information services, elderly-friendly route guidance, personal mobility services, mobile devices, biometric authentication interface and multi-modal dialogue mitigation and other smart workplace applications). Applications are all integrated as a unified, dynamic service batch, managed by the OASIS Service Centre and supporting all types of mobile devices (tablet PC, PDA, smartphone, automotive device, ITV, infokiosk) and all types of environments (living labs, sheltered homes, private homes, two car demonstrators, public transport, DSRT, etc.) in 4 Pilot sites Europewide.

As user friendliness and acceptability is a top priority for the project, a user-centred-design approach is followed along the service and application development. Tested iteratively and thoroughly by hundreds of end users, their caregivers and other stakeholders, the OASIS platform and applications will be optimised and submitted for standardization by the purpose-established OASIS worldwide Industrial Forum."

10.4.4.7 SHAPE

Title: Semantically-enabled heterogeneous service architecture and platforms engineering
Research area: ICT-2007.1.2 Service and Software Architectures, Infrastructures and Engineering
Project start date: [2007-12-01]
Semicolon Relevance: Possibly a bit on the technical side wrt Semicolon. Nevertheless: Semantics, Metamodel, Model based, interoperability by resolution (accepting/supporting heterogeneity)

Description
"The objective of SHAPE is to support the development and realization of enterprise systems based on a Semantically-enabled Heterogeneous Service Architecture (SHA). SHA extends service-oriented architecture (SOA) with semantics and heterogeneous infrastructures (Web services, agents, Semantic Web Services, P2P and grid) under a unified service-oriented approach.

To achieve this, SHAPE will develop a model-driven engineering (MDE) tool-supported methodology. SHAPE will take an active role in the standardisation of metamodels and languages for SHA. The technical results will be compliant with the proposed standards to ensure high industry acceptance.

In current SOA approaches, business requirements and technical details are intertwined constraining the evolution of service-oriented business solutions. SHAPE will provide appropriate metamodels and languages, methods and tools to separate the different viewpoints of SOA for the development of semantically enabled, flexible and adaptive business services on a rich SHA infrastructure. SHAPE addresses the target outcomes of the objectives in ICT-2007.1.2.

The work in SHAPE will be organised in seven work packages: Industrial use cases, Model-driven methodology and architecture, Metamodels and languages, Modelling tools and services, Model transformations and deployment, Standardisation, dissemination and exploitation, and Project management."

10.4.4.8 SYNERGY

Title: Supporting highly adaptive Network enterprise collaboration through semantically enabled knowledge services
Research area: ICT-2007.1.3 ICT in support of the networked enterprise
Project start date: [2008-02-01]
Semicolon Relevance: Interoperability and self-adaptation, semantically enabled knowledge services

Description
The next phase of enterprise interoperability is the sharing of knowledge within a Virtual Organisation (VO) to the mutual benefit of all VO partners. Such knowledge will be a driver for new enhanced collaborative enterprises, able to achieve the global visions of enterprise interoperability. The SYNERGY project envisages the delivery of Collaboration Knowledge services through trusted third parties offering web-based, pay on demand services, exploitable through interoperability service utilities (ISUs).

The overall aim of SYNERGY is to enhance support of the networked enterprise in the successful, timely creation of, and participation in collaborative VOs by providing an infrastructure and services to discover, capture, deliver and apply knowledge relevant to collaboration creation and operation.

Specifically SYNERGY aims to:
(a) provide semantic ontology-based modelling of knowledge structures on collaborative working;
(b) develop the service-oriented self-adaptive SYNERGY holistic solution for knowledge-based collaboration services; and
(c) facilitate the testing and evaluation of the efficiency and effectiveness of the SYNERGY solution in concrete case studies.

10.4.4.9 TARGET

Title: Transformative, adaptive, responsive and engaging Environment
Research area: ICT-2007.4.3 Digital libraries and technology-enhanced learning (ICT-2007.4.3)
Project start date: [2009-01-01]
Semicolon Relevance: Breaking down barriers of interoperability through better understanding, - fostered by game.

Description
"There is a need to reduce the knowledge worker "time-to-competence". The main aim is to develop a new genre of TEL environment that supports rapid competence development within the domains of innovation and project management.

The TARGET environment consists of a learning process supported by the TARGET platform, consisting of a main core of a serious game. Here, the learner is presented with complex situations in the form of game scenarios. Interacting with the game results in experiences that are gradually honed into knowledge."

10.4.4.10 4WARD

Title: 4WARD - Architecture and design for the future Internet

Research area: ICT-2007.1.1 The network of the future
Project start date: [2008-01-01]
Semicolon Relevance: Federated architecture that is "self-managing" with robustness that is "leveraging diversity". Although perhaps not intended for organizations, - similar requirements are being explored by Semicolon through the Name Message Case for actors in a federated process management system.

Description
"Today's network architectures are stifling innovation, restricting it mostly to the application level while the need for structural change is increasingly evident. The absence of adequate facilities to design, optimise and interoperate new networks currently forces a convergence to an architecture that is suboptimal for many applications and that cannot support innovations within itself, the Internet.

4WARD overcomes this impasse through a set of radical architectural approaches built on our strong mobile and wireless background.. We improve our ability to design inter-operable and complementary families of network architectures. We enable the co-existence of multiple networks on common platforms through carrier-grade virtualisation for networking resources. We enhance the utility of networks by making them self-managing. We increase their robustness and efficiency by leveraging diversity. Finally we improve application support by a new information-centric paradigm in place of the old host-centric approach. These solutions will embrace the full range of technologies, from fibre backbones to wireless and sensor networks.

The 4WARD results will allow new markets to appear, redefining business roles and creating new economic models. We will establish the Future Internet Forum as a leading standards body, enabling these new markets and opening them for old and new players alike, increasing opportunities for competition and cooperation and creating new products and services."

11 REFERENCES PART I

[AthenaDA4.2] Specification of interoperability framework and profiles, guidelines and best practices Deliverable A4.2 from the Athena Project
http://interop-vlab.eu/ei_public_deliverables/athena-deliverables/A4/d-a4.2/

[Bekkers, 2007] V. Bekkers: The government of back-office integration. Organizing co-operation between information domains. *Public Management Review, 9*(3), 377-400.

[BPMN] See [OMG BPMN]

[Eynon-Margetts, 2007] Eynon-Margetts:Organisational solution for overcoming barriers to eGovernment, European Journalof ePractice

[Concept research programme] http://www.concept.ntnu.no/index_engelsk.htm

[Cordis EI Roadmap, 2006] Enterprise Interoperability, Research Roadmap. 31 July 2006. Version 4.0.
http://cordis.europa.eu/ist/ict-ent-net/ei-roadmap_en.htm

[FAOS-report, 2008] Common ICT-Architecture in Public Sector (in Norwegian).
http://www.regjeringen.no/upload/FAD/Vedlegg/IKT-politikk/Felles_IKT_arkitektur_off_sektor.pdf

[eGov plan, 2003] Strategy for ICT in the Public Sector 2003 – 2005.
http://www.regjeringen.no/upload/kilde/aad/prm/2003/0002/ddd/pdfv/183023-ikt_engelsk_-endelig2.pdf

[eGov plan, 2005] eNorway 2005 (in Norwegian: eNorge 2005)
http://www.regjeringen.no/upload/kilde/mod/red/2000/0002/ddd/pdfv/154031-enorge_2005_endelig.pdf

[eGov plan, 2009] eNorway 2009, The Digital Leap (in Norwegian: eNorge 2009 – det digitale spranget)
http://www.regjeringen.no/upload/FAD/Vedlegg/IKT-politikk/enorge_2009_komplett.pdf

[Health plan, 1997] More Health for every BIT, Information Technology for a better Health Service, action plan 1997-2000 (in Norwegian: Mer Helse for hver B *IT*, Informasjonsteknologi for en bedre helsetjeneste, handlingsplan 1997-2000)
http://www.regjeringen.no/nb/dep/hod/dok/veiledninger_og_brosjyrer/1996/Mer-helse-for-hver-bIT.html?id=87401

[Health plan, 2001] Electronic collaboration in the health and social sector, action plan 2001-2003 (in Norwegian: "Si @", Statlig tiltaksplan 2001-2003: Elektronisk samhandling i helse- og sosialsektoren)
http://www.regjeringen.no/nb/dep/hod/dok/veiledninger_og_brosjyrer/2001/I-1018-B.html?id=87689

[Health plan, 2004] Te@mwork 2007, Electronic Cooperation in the Health and Social Sector, National Strategy 2004-2007 for Norway.
http://www.shdir.no/vp/multimedia/archive/00010/IS-1267_E_10176a.pdf

[Health plan, 2008] Teamwork 2.0, National strategy for electronic collaboration in the health and social sector 2008-2013 (in Norwegian: Samspill 2.0. Nasjonal strategi for elektronisk samhandling i helse og omsorgssektoren)
http://www.shdir.no/vp/multimedia/archive/00047/Samspill_2_0_-strate_47719a.pdf

[IDABC EIF 2.0] Draft document as basis for EIF 2.0. IDABC.
http://ec.europa.eu/idabc/servlets/Doc?id=31597

[INTEROP-VLab] INTEROP-VLab: The European Virtual Laboratory for Enterprise interoperability
http://interop-vlab.eu/interop-vlab-network/INTEROP-VLab/

[ISO 2004]: ISO/IEC 11179, Information Technology – Metadata registries (MDR), http://metadata-stds.org/11179/.

[OMG] Object Management Group
(OMG is an international, open membership, not-for-profit computer industry consortium.)
http://www.omg.org/

[OMG BMM] OMGs Business Motivation Model – Models for Business Plans
http://www.omg.org/technology/documents/bms_spec_catalog.htm

[OMG BPMN] OMGs Business Process Modeling Notation
http://www.bpmn.org/

[OMG CWM] OMGs Common Warehouse Model
http://www.omg.org/CWM

[Report no 17, 2006] . 'An Information Society for All', Report no. 17 (2006-2007) to the Norwegian parliament (in Norwegian).
http://www.regjeringen.no/Rpub/STM/20062007/017/PDFS/STM200620070017000DDDPDFS.pdf

[Riksrev Doc 3:7, 2008] The Office of the Auditor General in Norway inquiry about ICT in hospitals and electronic collaboration in the health sector. (in Norwegian: Riksrevisjonens undersøkelse om IKT i sykehus og elektronisk samhandling i helsetjenesten, Dokument nr. 3:7 (2007-2008)). Published 22nd April 2008.
http://www.riksrevisjonen.no/NR/rdonlyres/100CD1C0-9B8B-40BB-B35C-82D10FF80F3A/0/Dok_3_7_2007_2008.pdf

[Riksrev Doc 3:12, 2008] The Office of the Auditor General in Norway inquiry about electronic information exchange and service development in public sector. (in Norwegian: Riksrevisjonens undersøkelse av elektronisk informasjonsutveksling og tjenesteutvikling i offentlig sektor, Dokument nr. 3:12 (2007-2008)). Published 1st July 2008.
http://www.riksrevisjonen.no/NR/rdonlyres/323D98A4-BEE9-4E49-BDC7-4879337C3B39/0/Dok_3_12_2007_2008.pdf

[SemanticGov, Extended Project Presentation, 2006] V. Peristeras, T. Vitvar, A. v. Overeem: SemanticGov: D1.1: Extended Project Presentation.
http://www.semantic-gov.org/index.php?name=Deliverables#

[SemanticGov, Conceptual Analyses] R. Baldoni et al: WP01 - Overall Conceptual Analyses, D1.2 – State of the Art Report, Revision 19
http://www.semantic-gov.org/index.php?name=Deliverables#

[Smith-Meyer, 2005] H.Smith-Meyer, F.Lillehagen, H.D.Jørgensen, D.Karlsen, E.Dehli, S.Murberg, H.G.Solheim : Using Models and Model-based Transformation to Generate Operational Context-sensitive Workplaces. *Proceedings Collaborative Engineering 2005.*

[UN eGov survey, 2008] UN e-Government survey 2008, From E-government to Connected Governance
http://unpan1.un.org/intradoc/groups/public/documents/UN/UNPAN028607.pdf

[Updates on project governance in Norway] Presentation by Peder Berg, Deputy Director General, Ministry of Finance.
References

Part II

Organisational interoperability

12 ORGANISATIONAL INTEROPERABILITY

12.1 Organisational interoperability [some definitions]:

"Organisational interoperability deals with modelling organisational processes, aligning information architectures with organisational goals, and helping these processes to co-operate." (Manuela Finetti, Head of IDA Unit, European Commision)

"Organisational interoperability is concerned with defining business goals, modelling business processes and bringing about the collaboration of administrations that wish to exchange information and may have different internal structures and processes. Moreover, organisational interoperability aims at addressing the requirements of the user community by making services available, easily identifiable, accessible and user-oriented." (IDABC used this definition in 2005)

"Organisational interoperability: The processes by which different organisations such as different public administrations collaborate to achieve their mutually beneficial, mutually agreed eGovernment service-related goals." (IDABC, EIF draft 2)

12.2 Executive summary on Part II - Organisational interoperability

The Semicolon project addresses semantic and organisational interoperability. This part (part II) focuses on the organisational aspects of interoperability and refers to:

- Organisation science
- Enterprise communication
- Enterprise modelling
- Business process management (***not*** in focus in this version)
- Economical theory (Transaction cost theory, Agency theory, Alliance theory, Network theory, Contractual theory, Theory of competencies, Relational exchange theory, Stakeholder theory, Theory of organisational boundaries, Social exchange theory, Institutional theory, Organisational climate).
- Information governance (Data, Information Security, Information Technology, Internet)

These areas are described in the following and are both explained in further details and also given references to theory and relevant articles/sources. Chapter 14 gives the state of the art references within these given areas, and in Chapter 15 we refer to some best practice examples within the Information governance part. An own literature review based on the economical theory part is done in Appendix 1 to this report.

13 STATE OF THE ART

13.1 Interoperability and Organisation Science

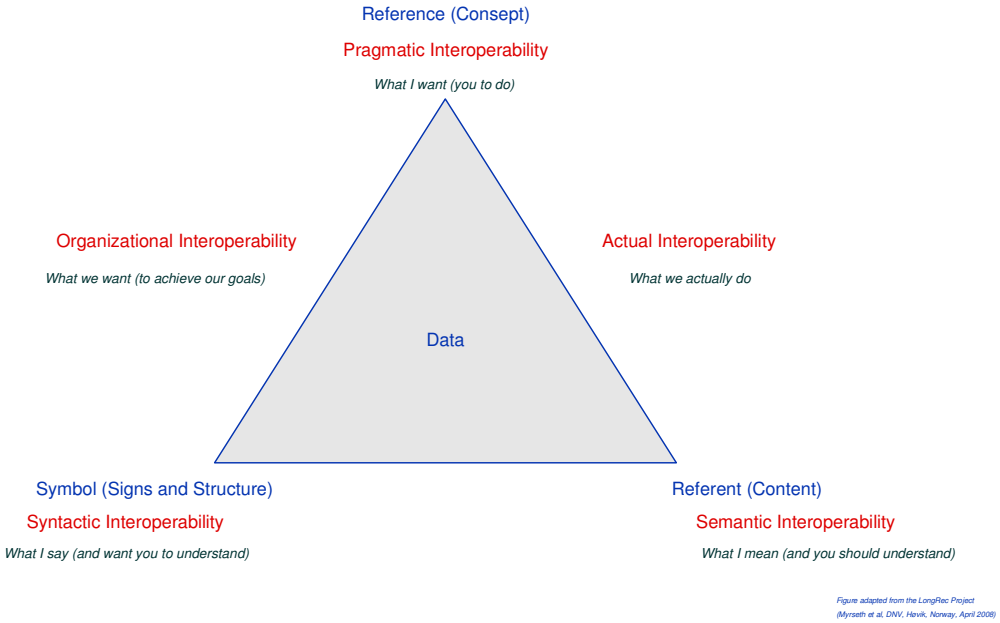


Figure: The semiotic triangle (see also chapter **18.1 "**Some clarifications about semantics")

In order to relate the above model with the realities of human enterprise we invoke concepts from organisation science [Thompson]. Specifically, we use ideas from the structural contingency theory [Galbraith] and information processing in organisations [Scott] to introduce the concepts of complexity[5], uncertainty[6] ambiguity[7] , and the resulting interdependence[8]

- *Complexity* is a measure of the number of (partial) solutions that are required to adequately address problems facing the organisation.
- *Uncertainty* is a measure of the difference between available and required information to address problems facing the organisation
- *Ambiguity* is a measure of the number of interchangeable problem interpretations that are relevant for a given organisational solution.
- *Interdependence* is a measure of how actors in the organisation depend on each other in order to carry out the tasks for which they are responsible.

[5] http://en.wikipedia.org/wiki/Complexity

[6] http://en.wikipedia.org/wiki/Uncertainty

[7] http://en.wikipedia.org/wiki/Ambiguity

[8] http://en.wikipedia.org/wiki/Interdependence

These measures can be operationalized and measured using Quality Function Deployment (QFD)[9] and Design Structure Matrix (DSM)[10] tools from Project Management[11], and operationalized and implemented for simulation of organisation [Christiansen].

Pulling together the above issues from of information theory, semiotics and organisation science we illustrate a combined Information Quality Framework –

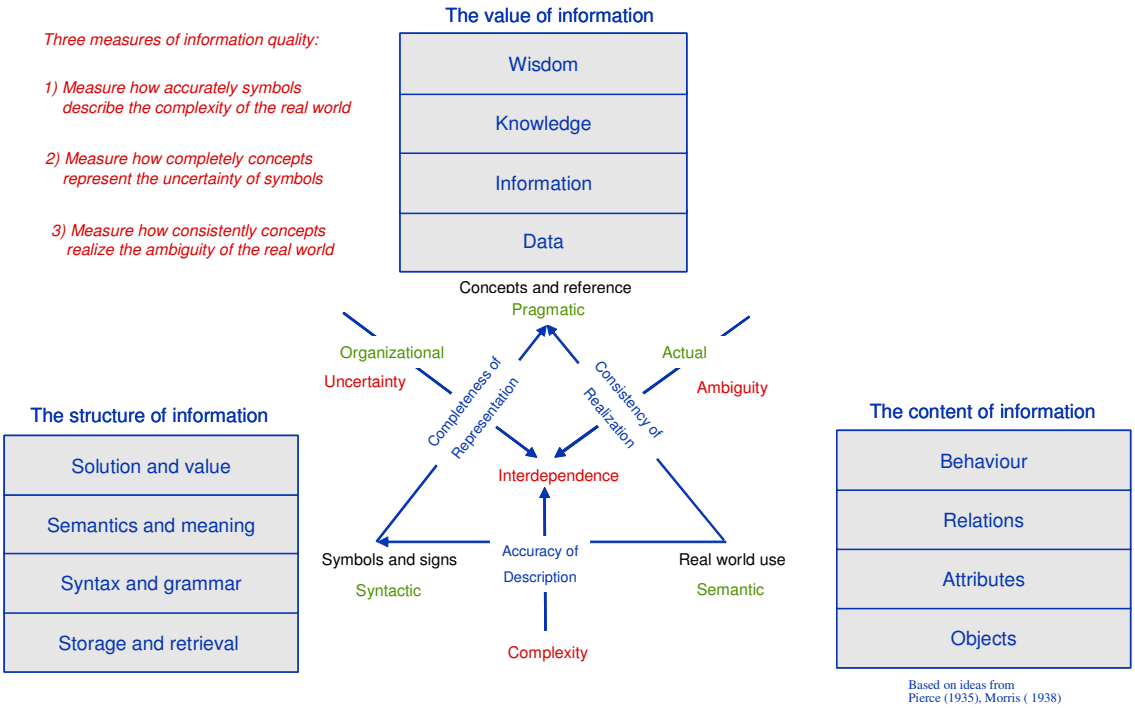


Figure 5: A combined Information Quality Framework

This Information Quality Framework needs further concretization, operationalization, implementation and testing before it can serve as a basis for practical tools that can be used to deliver (first) descriptive and (then eventually) prescriptive Information Quality Solutions to customers.

13.2 Enterprise Communication, Enterprise Modeling

Enterprise Modeling (EM) provides means to communicate and understand what an enterprise does, what it can do, and why, where and how enterprises should cooperate/interoperate for the mutual and

[9] http://en.wikipedia.org/wiki/Quality_function_deployment

[10] http://en.wikipedia.org/wiki/Design_Structure_Matrix

[11] http://en.wikipedia.org/wiki/Project_management

collective benefit to themselves and their "customers". ISO 19439 (2003) defines an Enterprise model to be"a representation of what an enterprise intends to accomplish and how it operates."

Although there are examples and research [Smith-Meyer 2005] of means and methods [OMG MDD] to operationalize the models produced by Enterprise modeling, - its first and foremost angle of value is arguably communication, at least at this time. For the enterprise model to become a medium of interest, however, - the model needs to be consistent with the way the enterprise actually operates, either by reflecting it correctly, - or by dictating it. This usually implies mappings across different abstraction levels.

Typically, an Enterprise Model makes some of the following aspects of an enterprise explicit:

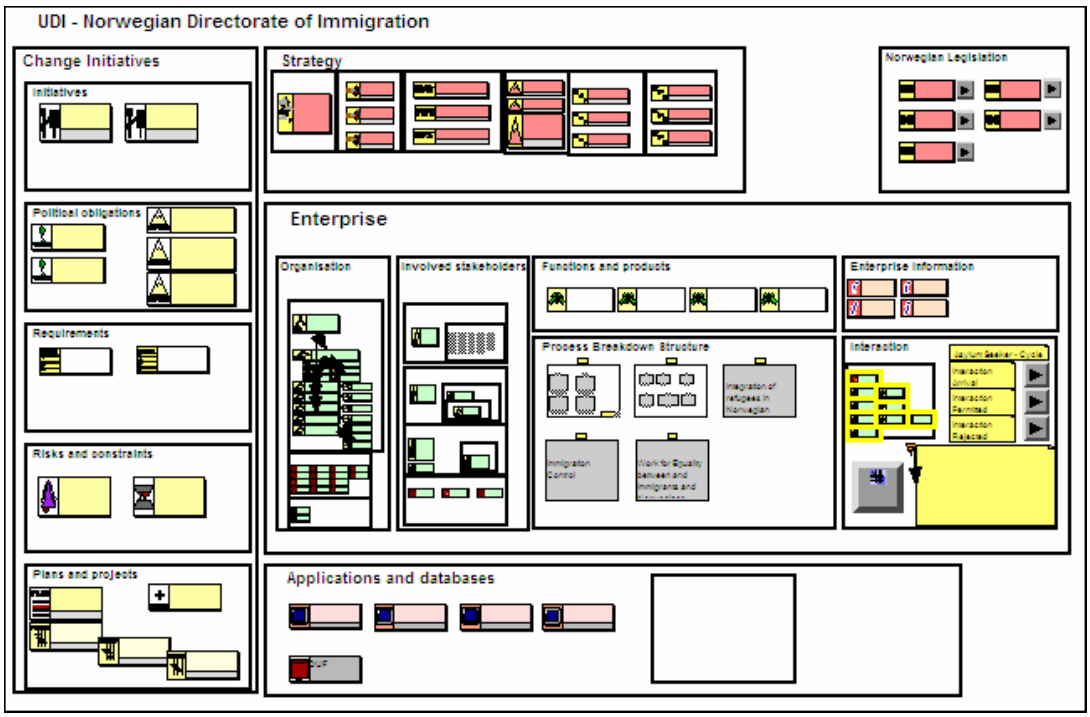
- products and services – what the enterprise delivers
- business processes – how products and services are produced and delivered
- Organization, People, Roles and Responsibilities, Locations
- Strategies and Goals, - Vision, Mission and Values,
  Key Performance indicators, Strengths and Weaknesses, Threats and Opportunities,
  Challenges, Risk and Critical points
- Measures, Key performance Indicators
- Information and Knowledge – Core Competences, Master Data, …
- Law and regulations
- Systems, Technology, Standards used or recommended
- The interconnections/relations of all the aspects above
- And last but not least: Change and Transformation plans, Possibilities, and
- (History and Traceability is usually not "high-on-the-list", - if at all considered)
- (Metadata about the maturity, precision, uncertainty etc. of model elements, - possibly carrying the views of different actors

Enterprise Modeling is often associated with graphical modeling tools, since a graphical presentation is thought to offer a pedagogical element of communication, - following the accepted idea that "a picture tells more than a thousand words".

Note that requirements to correctness and precision can be relaxed in order to facilitate communication. It should also find its most useful level of detail and abstraction in order to facilitate its objective: a common understanding among participating actors and interested parties. A too detailed level will loose most participants, and a too high level can be out of touch with reality and the work people are actually doing.

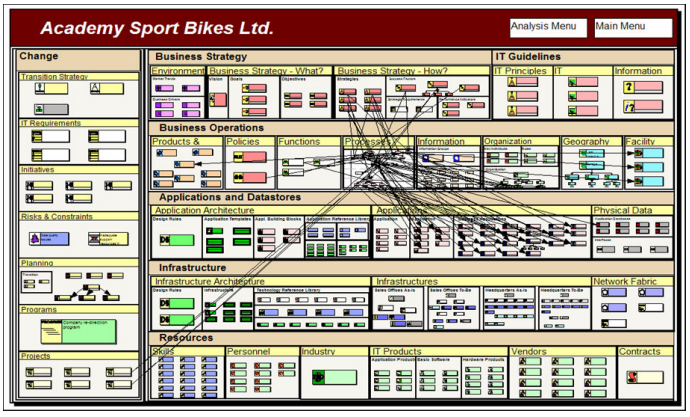
The following figure shows an example "bird view" of an enterprise model in one tool[12]:

[12] Figure from Athena A1 SoA, figure 44, page 98 (http://interop-vlab.eu/ei_public_deliverables/athena-deliverables/A1/d-a1-1.1/)



(**Figure XX** showing a graphical layout of multiple enterprise domains. The figure below includes some relationships from members of one aspect to members of other aspects)

Seen from Norwegian Government, we note that OppgaveRegisteret (http://www.brreg.no/registrene/oppgave/) certainly constitutes a partial enterprise model of Norwegian Government. So is Lovdata (www.lovdata.no), which is a listing of Norwegian Law. We note that it both centrally maintained, and updated through satellites communication with a central body that edits and provide central control of quality. A key to Enterprise modeling is however how the various aspects are interconnected, and one of the challenges is how to distribute responsibility for maintaining and communicating knowledge about the relations.



We doubt that the current approach is a sustainable for interoperability among governmental bodies, at least not for the full range of characteristics and enterprise data described above. Oppgaveregisteret should retain its function and methods as is, - in order to provide a trusted and legitimate source of information. It is however also necessary to provide an "distributed and collective enterprise communication platform" where each actor can independently describe themselves the way they want external partner to see them, - perhaps with common interoperation descriptions of meeting places, services offered and utilized, shared events and data, - etc. Even if each actor can and should "tell the truth about himself" it might be useful also for other actors to describe their view of others business

than their own, where the analysis and syncronization of such descriptions is useful for the process of increasing interoperability, and interoperation where this implies increased utility.

How can such a "distributed and collective enterprise communication platform" be realized? The major requirements seem to be:

- It must be distributed, - independently editable and maintainable by each of the now 700 governmental bodies (and perhaps a number of other private actors or interest groups that provide services on behalf of government).
- It must be collective, - each of the distributed entries must be interconnected to provide a collective and holistic picture that everyone can understand, utilize and act in relation to
- It must be available, - in general, all information about the whats, whys, and hows should be available to all. There are exceptions: Some "hows" need to be protected and only available to clusters of actors (how are tax statements checked), - and some "whats" (the focus of secret services) will remain out of the picture. The exceptions should simply be excluded from the picture, - until a satisfactory platform is reached for the elements that should be publicly available.
  Availability implies both reading and modification, with the ease of use and flexibility that empowers any business responsible to utilize it.
- It must be understandable, - utilizing defined languages and with pedagogical means. Graphical modeling is a popular means to communicate and transfer knowledge. It is however not clear how to make use of graphical editing tools (with manual layout) in environments where the model is built from so many different units, with so many different people and units involved. Furthermore, manually designed graphical diagrams are brittle with regards to change, - but the manual design is important for pedagogical quality. Automated on-the-fly layout algorithms, perhaps guided semi automatically or automatically by manually designed patterns, may increase understanding without harming the required agility.
- It must be agile and traceable, - and should never be expected to be complete , precise or accurate.
- It must however have the security needed to log who does what correctly, probably together with other functionality for establishing trust.
- It must be analyzable, both algorithmically and manually. Algorithms can provide measures of interoperation and interoperability: of connectivity and interoperation, vertically and horizontically, between government bodies and the exposure to external parties ("customers" or "partners"), of shared services and processes, shared data and metadata, utilization of data acquired from other actors, - together with measures of change from year to year
- Finally, it should lay the groundwork for a link to operation, with means to measure operational data, - not only the descriptive data of enterprise models. This implies some link, but we note that the descriptive and operational models in questions might be at different abstraction layers, and therefore not tightly coupled

To some degree, the platform needs many of the characteristics of the Web1.0 internet platform that made "the man on the street" capable of publishing whatever he liked to the rest of the world. But

there are new requirements to form, content and analyzability that takes it a bit further, - even without moving into operation.

The following development areas offer tools and functionality that could give a base to such a platform:

- Enterprise Modeling for Collaborative Engineering. This was the focus in the Athena project and the Athena Interoperability Framework that we described in Part I above. Advanced modeling tools as such is perhaps more in focus than what is required for the platform we describe above, and later experiences of Web2.0 "collaboration" are not explored. Also, there is perhaps more focus on a collaborative design of a common model than the interconnection of multiple more independent models. Nevertheless, such differences might not be substantial and methods and tools developed for collaborative engineering might very well be useful in the Semicolon context.
- Enterprise architecture (EA). A number of frameworks exist to define a common architecture and choice of methodology, modeling purpose, abstraction level, languages, etc. in order to support development projects to build components that will be interoperable with other components developed within the same EA or EA framework[13] (such as Zachman, TOGAF, FEAF, OIO[14], DODAF, …)
- The workflow and business process management movement of process orientation, represented by Workflow Management Coalition (http://www.wfmc.org/) and BPMI/OMGs BPMN
- OMGs Model Driven Architecture (MDA) offer methods and means for software production
- Micosoft has provided the world with interoperability through their Office Platform. We might not like the dominance of one vendor in any area, but any extension of the interoperation-enabling brought to us by Microsoft will have potential to improve interoperability simply by expected market adoption.
- The Social Web, Web2.0 (and announced "3.0"). Facebook, LinkedIn, Freebase offer platforms where people can be said to be interoperable based in a common model that they populate in a collective and distributed manner. The three former also offer other actors to build extension based on whatever theme one might find useful. Developers have taken on the challenge, and the drivers for extensions seem to be present. This is however relatively "young", and how such patterns of functionality could be utilized within the government sector is open to debate and exploration.
- W3C and the Semantic Web build on the distributedness of Web1.0 to include common semantics and distributed models across the net. Concepts like SIOC "Semantically Interlinked Online Communities" and "Linking Open Data"[15] also offer operational platforms for sharing of models in a distributed environment. . Security and Trust issues are perhaps most important when we think of these technologies to be relatively immature with regards to Semicolon.

[13] For a list of some dominate frameworks, see English Wikipedia: http://en.wikipedia.org/wiki/Enterprise_Architecture_framework

[14] OIO – Denmarks "Offentlig (Public) Informasjon Online" http://digitaliser.dk/Default.aspx

[15] http://esw.w3.org/topic/SweoIG/TaskForces/CommunityProjects/LinkingOpenData

The semantic web and Social We2.0 sites offer architecturally different models of overlapping functionality. Both can enable distributed working environments, but for the semantic web, the data is also distributed around the world, - like on the Web1.0. We note that the centralized model of the Web2.0 sites perhaps offer better control and means of lifting or promoting models.

13.2.1 Gartner's view on enterprise architecture tools and their requirements

Gartners report on "criteria to select an enterprise architecture tool" (Nov 2007) gives some insight to the commercial state of affairs:

Key Findings

• Selection of the most appropriate EA tool is rarely clear-cut. In most cases, a range of criteria need to be evaluated and weighed.
• There are five key aspects of an EA tool's functionality that should be considered:
• How the tool organizes information
• How it presents the information to support the needs of various stakeholder groups
• How it is able to analyze the information to assist EA
• The extent to which it is able to exchange information with other tools and/or facilitate the entry of information directly into the tool
• The administrative support that the tool provides
• Other criteria, such as the cost of the tool and the viability of the vendor, should also be considered.

Recommendations

• Base your EA tool evaluation criteria on those that are described here, but be prepared to supplement or modify them to address your unique situation.
• When evaluating EA tools, weight the criteria to reflect their relative importance to your organisation.

[…]

Other points related to structuring architecture-related information include:
• Support for links to documents and other objects stored outside the repository
• The power and ease of use of any metamodel customization capabilities
• Support for architecture frameworks such as the Zachman Framework and The Open Group Architecture Framework (TOGAF), as well as the ability to customize such frameworks or add new ones
• Support for ontologies (ontologies are defined in "Sharing Semantics Across Applications")

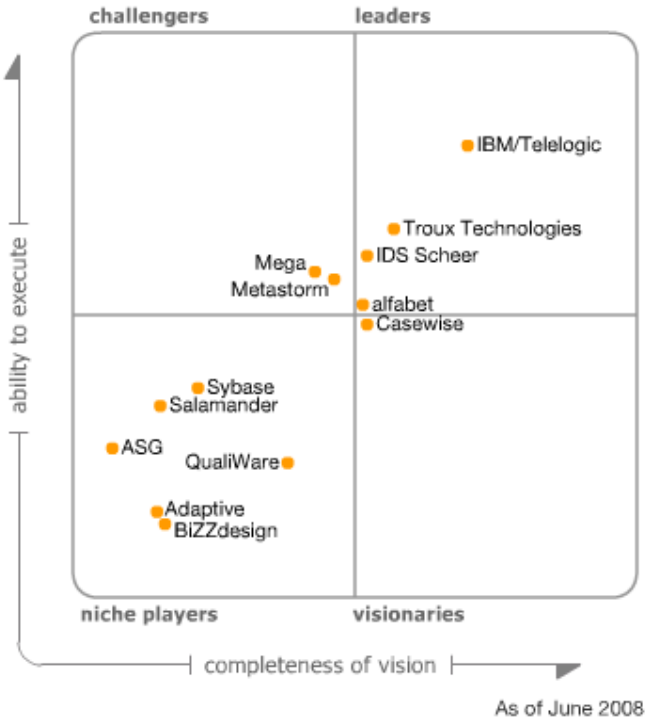
Gartners analysis reflects an immature market where the vendors differ in history/root, understanding, market angle and focus. Even if common requirements exist, for example with the support for standards and modeling languages, the vendors offer different solutions, and depend on the customers to customize according to their needs.

The recommendation to focus on customer/user need reflects a wide variety in approach and context. No vendor can offer a general solutions that has all the answers, - with mature and proven functionality for a positive ROI.

Each customer is therefore obliged to have a good idea of what he/she wants to achieve and to carefully design their own approach and solution. Even if frameworks exist, the variety within them can be considerable.

It is worth noting that companies strive for uniqueness in their quest for competitive advantage, - and that functionality for customization therefore will remain important for enterprise modeling.

Nevertheless, Gartner publishes a report each year where the vendors are classified along the axis of ability to execute and leaders/visionaries, latest in June 2008:



About the three leaders it is worth adding that he IBM Telelogic builds on Popkins System Architect after acquisition, Troux Technologies builds on the Metis toolset after acquisition and IDS Sheer sells the ARIS Toolset.

For Semicolon and the Norwegian Government, such tools might form a basis both for common understanding and for measures of operation and quality. It is however not so clear how they should be utilized nor by whom in what roles, - in order to satisfy the needs as described above. To stay in front, - all tools will have to improve their collaborative and distributed functionalities, possibly guided by the social network platforms on the internet.

13.2.2 EM Pitfalls and Challenges:

13.2.2.1 Enterprise Modeling antipatterns

There are many ways to go wrong with enterprise modeling.
http://www.agilemodeling.com/essays/enterpriseModelingAntiPatterns.htm lists a number of issues.

In sum, *they reflect patterns where modeling is disconnected from the daily business of the firm.* Arguably, there is something wrong with the context of modeling work. One problem is that it often is a long path from modeling what to do and doing it, - with or without the support of the model. Not only in time, but the path can travel through different people, losing semantics of both intension and content, and over the barriers of different perspectives, like through the knowledge acquisition bottleneck of software and knowledge engineering.

Web2.0 patterns and social networks give rise to early ideas of some new possibilities, - perhaps in relation with vendors rising attention to team modeling functionality and model repository solutions. We have however, not identified results or literature in this area.

13.2.3 Business Process Management and Service Orientation

At this time it is not clear what role BPM will play in Semicolon with regards to interoperability. We expect that BPM can and will play a larger role in version 2 of this State of the Art document. The following is just *a brief commentary to touch on some of the issues*.

Business Process Management is an area where many disciplines meet. Process Orientation is, often in conjunction with Service Orientation, becoming increasingly popular for braking down the siloes of enterprises and establishing horizontal interoperation, but there is still no operational, available and general-use platform for simple execution support and tracking of a work process. Even if the encapsulation of legacy systems into Services that can be utilized in process steps offer a bridge to BPM, and Service Orientation can be realized by Business Processes the concepts and experiences have not produced a breakthrough approach with the agility and ease of use solutions that business people can utilize. There might be several reasons for this:

- A too strong focus on producing operational systems directly from strategic enterprise models, following the MDA approach of OMG, with insufficient speed in the turnaround from change in enterprise models to reflections in operational systems.
- Missing focus on the fact that not all processes are supported by Software Components. This is gradually changing, and human-oriented BPM is developing.
- A focus on automation and orchestration, - and missing support for human knowledge and interruptions. We can ask whether the hype of operational languages like BPEL (Business Process Execution Language) has become a setback with regards to supporting the business tasks that can be ad-hoc and human knowledge intensive. BPEL4PEOPLE quickly emerged to address the point, but without providing a useable platform.
- A dependency on software development processes, which ruins the quest for a quick feedback and realization cycle, by imposing financial needs, communication barriers (the software developer), transformation of needs into implementation languages (and thereby making sure the business people have lost track and control of what is going on), and time (making sure business people have lost interest and/or moved to other issues by the time the need for business change meets its realization)

The above issues, and more, stands in the way of business people being able to describe what they want to do and to describe how to do it, - in a way that they can get support from an operational system with the possibility of offering distributed tracking of affairs (how many cases where closed today? How many do we need to close within the next week in order to comply with rules, regulations and our own quality program?) and operational tracking of KPIs that motivate the desired behavior.

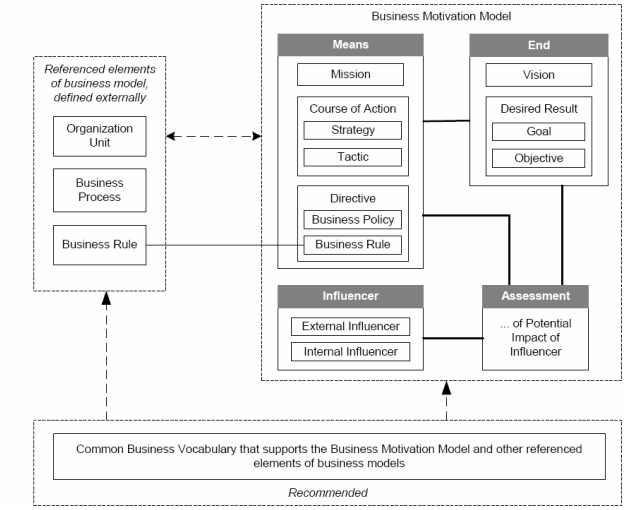
It is very possible that there are technological possibilities that will work and satisfy the needs of Process Oriented Business People with considerable value to such organization. Possibly, the proposed solutions are too complex, so complex that the users do not collectively understand their role and possibilities of utilization.

Another challenge is the number of systems that users are exposed to, where many of these attempt to take control of the process. Thereby, we get many Business Process Management Systems, and less holistic business process management. We should not be surprised, since it is only natural that every function and tool will see themselves as the center.

BPMI (www.bpmi.org) rose in the early 2000s with a mission to make Business Process Modelling available to business users through a new Business Process Modeling language BPMN – Business Process Modeling Notation (www.omg.org/bpmn). Even if the "movement" appeared as a reaction to a software dependence, it merged with the more established OMG in 2005, produced early BPMN1.1 in 2008, - and BPMN2.0 is "in the process". With software-oriented OMGs focus on software development, Object Orientation and unification of languages, it will be interesting to see if BPMIs original quest for making process management available to business people is still in the drivers seat, - or not.

Gartner has kept the quest for user-oriented BPM in their Business Process Management Suite (BPMS) Magic Quadrant Report. The Suite is intended to be a complete set of tools for giving operational support to knowledge and process models that can (preferably) be maintained by business users.

Another OMG product relevant to the subject of Semicolon is the Business Motivation Model BMM, addressing some of the aspects described in our chapter on enterprise modeling above.



(**Figure XX** from http://www.omg.org/spec/BMM/1.0/, "Overview of BMM")

A full description of the state of the art of operational platforms for BPM would include the products associated with WfMC (Workflow Management Coalision) and Gartners classification of BPMSs (Business Process Management Suites). In addition (and not reflected in the Gartner BPMS evaluation), the market power of Microsoft, Oracle and IBM forces their relevance: Microsoft has announced a new platform "*Oslo*" with the explicit reference language "M2", Oracle has aquired middleware platforms and utilizes the Enterprise modeling platform ARIS for their Business Process Automation platform, and IBM has aquired Popkins System Architect to be possible to integrate in their WebSphare platform.

The Microsoft Platform is also relevant for Semicolon due to the fact that Accentures accepted offer for the realization of AltInn2 is based on the Microsoft Platform.

13.3 Economical theory

13.3.1 The concept of e-government

The term “electronic government” or “e-government” or "digital government" appeared about a decade ago, and there is no commonly accepted definition. Some see e-government as the migration of government information and services to an on-line delivery mode, where the scope of e-government covers the interaction between government and citizens (G2C), government and business enterprises (G2B), and inter-agency dealing (G2G). Others see e-government as the provision of routine government information and transactions using electronic means, most notably those using Internet technology, whether delivery at home, at work, or through public kiosks.

It is an underlying assumption in this research that Internet technologies and specifically e-government should have as their main purpose the improvement of the ways in which government serves its citizens and the ways in which citizens interact with public institutions. This philosophy of e-government implies that for e-government to be anything more than automated service provision, it needs to reach far beyond the conduct of routine government business to embrace social, economic and political change.

Some stress that successful e-government programs should not only be based on the perceived efficiency gains for government itself, but rather on the satisfaction of consumers. For example, the UK’s “techno-centric model” has been criticized for failing to engage citizens as anticipated, underplaying the importance of knowledge management and clashing with traditional values of public service.

Some factors which seem important from the perspective of the suppliers of e-government include: the capacity for significant organisational change, the development of leadership skills, a grasp of the distinction between “hard” (technological factors) versus “soft” (human factors), and understanding of the differences in catering for the private and public sectors, and for citizens in developed and developing nations.

From the perspective of citizens’ needs (the “demand side” in economic terms), it might be stressed the vital role of factors such as the impact on citizens of transaction costs, an understanding of cultural barriers, for example social exclusion caused by the problem of unequal access to the Internet, citizens’ expectations of government services and their degrees of acceptance of technological innovations, and possible mismatches between governmental and social uses of the Internet.

For example, Vietnam is transforming into a networked society where more people are becoming connected, and more advanced applications, such as e-government, are becoming available. From 2000, the Government of Vietnam determined that, with Vietnam integrating more comprehensively into the global economy, the building of an effective e-government would help to facilitate its capacity to manage resources, implement sound policies and better satisfy the needs of citizens. Vietnam has a government official (an e-government champion) in charge of all e-government activities who liaises with other departments and ministries to ensure interoperability and interconnectivity (United Nations, 2008). Statistical data suggested that in May 2007 there were about 16 million Internet users, and 70 million others were living and working without the use of computers in Vietnam. This has to be taken into account when developing digital government in the nation.

E-government implies fundamental knowledge redistribution and requires a careful rethinking of the management of information resources and knowledge bases. E-government strategies include information dissemination on searchable databases, customer satisfaction, online translations,

implementing the use of e-checks and be able to measure government systems. It aims to offer accessibility to government information and services for citizens, businesses and government agencies thereby improving the quality of e-services and providing greater opportunities for participation in democratic institutions and processes.

Electronic government is the delivery of services to citizens via the Internet. The goal of e-government is to capture benefits of the electronic economy. Although there is sparse information about the quality and efficiency of e-government initiatives, an increasing number of governmental units are incorporating or expanding the use of information technologies into many of their activities (Esteves & Joseph, 2008).

Digital government has been considered a powerful strategy for administrative reform. However, projects looking for benefits to service quality or more effective and efficient government programs face a plethora of technical, organisational, and institutional challenges. As the organisational complexity of the projects increases and more agencies collaborate and share information both potential benefits and challenges increase (Luna-Reyes, Gil-Garcia, & Cruz, 2007).

The emerging IT-for-development approach towards public sector transformation is creating new perceptions about government and governance. The twin objective of achieving further improvements in service delivery and efficacy in government functioning is bringing about a rethinking of the role of IT. Governments are increasingly looking towards e-government-as-a-whole concept, which focuses on the provision of services at the front-end, supported by integration, consolidation and innovation in back-end processes and systems to achieve maximum cost savings and improved service delivery (United Nations, 2008).

Governments all over the world are recognizing e-government as a strategic option to fine-tune their internal and external operations. In order to foster citizen-centric services, they need to integrate themselves as well as stakeholders vertically and horizontally. This can be achieved by bringing the efficiencies and experiences of e-business to e-government. That requires new e-business models for government solutions that reduce cost and improve service effectiveness (Papazoglou & Ribbers, 2006).

Digital government, e-government, and e-governance are all terms that have become synonymous with the use of information and communication technologies in government agencies. Inter-organisational information integration has become a key enabler for e-government. Integrating and sharing information across traditional government boundaries involves complex interactions among and with technical and organisational processes. From a technical perspective, systems designers and developers must regularly overcome problems related to the existence of multiple platforms, diverse database designs and data structures, highly variable data quality, and incompatible network infrastructure. From an organisational perspective, these technical processes often involve new work processes, mobilization of limited resources, and evolving inter-organisational relationships. These necessary changes are influenced by specific types of social interaction, which take the form of group decision-making, learning, understanding, trust building, and conflict resolution (Pardo & Tayi, 2007).

E-government or digital government has become a global theme in governments pursuing an agenda of providing citizen services and increasing agency efficiency using IT. E-commerce and e-business is considered a trigger of e-government. Citizens have acquired competence in handling and operating e-commerce applications (net banking, e-ticketing, e-shopping) but most governments have not offered public digital self-service to the same degree and the public digital services that have been introduced are generally used to a lesser degree than private services. A government needs to consider markets for differentiated products and services to satisfy the variation of needs among its citizens. Electronic

markets may be just that kind of application that will trigger requisite information from citizens in order to offer them the relevant government services (Pedersen, Fountain, & Loukis, 2006).

Electronic markets for the allocation, financing and distribution of public goods were identified by Vragov and Kumar (2006). For example, electronic markets might be connecting citizens to pension reform (Ranerup, 2006). Salleh, Rohde and Green (2006) studied the effect of enacted capabilities on adoption of a government electronic procurement system by Malaysian SMEs, while Cui, Zhang, Zhang and Huan (2006) explored e-government impact on Shanghai organisations' informatization process.

13.3.2 Theoretical approaches

To understand the behavior of different government agencies at the level of organisational interoperability, some theories might help. In the following, we present a selection of theories, which might be applied to study organisational interoperability.

In our perspective of digital government, *transaction cost theory* has two implications. First, the organisation of government into agencies might be determined by transaction costs, where the number of agencies and the tasks of each agency should be such that transaction costs are minimized. Next, inter-organisational arrangements, such as information systems interoperability, should be such that transaction costs are minimized. Based on transaction cost theory, we might assume that when the degree of system interoperability is high between two cooperating agencies, then transaction costs are low. Opposite, when the degree of system interoperability is low between two cooperating agencies, then transaction costs are high. Therefore, benefits of interoperability can be measured in terms of transaction costs and changes in transaction costs over time.

Generally, organisations are hypothesized to choose organisational boundaries to minimize the sum of production and transaction cost (Williamson, 1979). Five attributes of business exchange are positively associated with transaction costs: (1) the necessity of investments in durable, specific assets; (2) infrequency of transacting; (3) task complexity and uncertainty; (4) difficulty in measuring task performance; and, (5) interdependencies with other transactions. The necessity of early investments in durable, transactions-specific assets (e.g. human and physical capital) shifts the balance of power between transaction participants, because in later renegotiations these costs are sunk costs of the party that incurs them. Infrequent transactions increase the likelihood of opportunistic behavior in later periods by reducing the threat of retribution. In situations where broader market reputations are at stake, infrequent transactions may be sustainable. However, even long-term contracts often do not provide sufficient adaptation mechanisms, and inflexibility may actually induce holdup. Task complexity, uncertainty and measurement problems exacerbate the problem of identifying and contracting for contingencies. Interdependencies introduce contingencies among transactions that suggest co-location (e.g. system-level sourcing) or that require high-level coordination (Anderson, Glenn, & Sedatole, 2000).

Transaction cost theory thus tells us that interoperability in digital government is influenced by (1) the need for specific hardware and software; (2) the frequency of information transacting; (3) task complexity and uncertainty; (4) difficulty in measuring task performance; and, (5) interdependencies with other transactions.

Agency theory has broadened the risk-sharing literature to include the agency problem that occurs when cooperating parties have different goals and division of labor. The cooperating parties are

engaged in an agency relationship defined as a contract under which one or more persons (the principal(s)) engage another person (agent) to perform some service on their behalf which involves delegating some decision making authority to the agent (Jensen & Meckling, 1976). Agency theory describes the relationship between the two parties using the metaphor of a contract. According to Eisenhardt (1985), agency theory is concerned with resolving two problems that can occur in agency relationships. The first is the agency problem that arises when the desires or goals of the principal and agent conflict and it is difficult or expensive for the principal to verify what the agent is actually doing. The second is the problem of risk sharing that arises when the principal and agent have different risk preferences. The first agency problem arises when the two parties do not share productivity gains. The risk-sharing problem might be the result of different attitudes towards the use of new technologies. Because the unit of analysis is the contract governing the relationship between the two parties, the focus of the theory is on determining the most efficient contract governing the principal-agent relationship given assumptions about people (e.g., self-interest, bounded rationality, risk aversion), organisations (e.g., goal conflict of members), and information (e.g., information is a commodity which can be purchased).

Alliance theory is concerned with partnership, often referred to as alliance. Das and Teng (2002b) studied how alliance conditions change over the different stages of alliance development to understand the development processes. They defined the following stages in the alliance development process:

- Formation Stage. Partner firms approach each other and negotiate the alliance. Partner firms then carry out the agreement and set up the alliance by committing various types of resources. The alliance is initiated and put into operation. Alliances will be formed only under certain conditions. These conditions include a relatively high level of collective strengths, a low level of inter-partner conflicts, and a high level of interdependencies.
- Operation Stage. Not only is the formation stage directly influenced by alliance conditions, the transition from the formation stage to the operation stage are also dictated by the same alliance conditions variables. During the operation stage, partner firms collaborate and implement all agreements of the alliance. The alliance will likely grow rapidly in size during this stage, somewhat akin to the growth stage of organisational life cycles. Other than the growth route, an alliance may also be reformed and/or terminated at this stage.
- Outcome Stage. During this stage, alliance performance becomes tangible and can, thus, be evaluated with some certainty. There are four possible outcomes for an alliance at this stage – stabilization, reformation, decline, and termination. A combination of outcomes is also possible, such as a termination after reformation. Alliance reformation and alliance termination do not necessarily signal alliance failure. Reformation and termination may be the best option under certain circumstances, such as the achievement of pre-set alliance objectives. Alliance condition variables continue to play a decisive role in the outcome stage. The particular alliance outcome will depend on the condition of the alliance.

Das and Teng (2003) discussed partner analysis and alliance performance. An important stream of research in the alliance literature is about partner selection. It emphasizes the desirability of a match between the partners, mainly in terms of their resource profiles. The approach is consistent with the resource-based theory of the firm, which suggests that competitors are defined by their resources profiles.

According to *network theory* (Afuah & Tucci, 2003), a network exhibits network externalities. An organisation exhibits network externalities when it becomes more valuable to members as more people

take advantage of it. A classic example from technology is the telephone, where the value for each subscriber increases exponentially with the number of network subscribers, to whom you can talk and get services from.

Contractual theory is concerned with the role of contracts in social systems. Luo (2002) examined how contract, cooperation, and performance are associated with one another. He argues that contract and cooperation are not substitutes but complements in relation to performance. A contract alone is insufficient to guide evolution and performance. Since organized crime often involves both intra-organisational as well as inter-organisational exchanges that become socially embedded over time, cooperation is an important safeguard mechanism mitigating external and internal hazards and overcoming adaptive limits of contracts. The simultaneous use of both contractual and cooperative mechanisms is particularly critical to organized crime in an uncertain environment. Relational contract theory was created by Macneil (2000), who has been doing relational contracts since the mid-1960s, and who by contract means relations among people who have exchanged, are exchanging, or expect to be exchanging in the future – in other words, exchange relations. He finds that experience has shown that the very idea of contract as relations in which exchange occurs – rather than as specific transactions, specific agreements, specific promises, specific exchanges, and the like – is extremely difficult for many people to grasp. Either that or they simply refuse to accept that contract can be defined as relations among people in an exchange. Macneil (2000) searched for roots to summarize contract in a useful manner. He tried to distill what he found into a manageable number of basic behavioral categories growing out of those roots. Since repeated human behavior invariably creates norms, these behavioral categories are also normative categories. He identified the following ten common contract behavioral patterns and norms: (1) Role integrity – requiring consistency, involving internal conflict, and being inherently complex, (2) Reciprocity – the principle of getting something back for something given, (3) Implementation of planning, (4) Effectuation of consent, (5) Flexibility, (6) Contractual solidarity, (7) The restitution, reliance, and expectation interests (the linking norms), (8) Creation and restraint of power (the power norm), (9) Proprietary of means, and (10) Harmonization with the social matrix, that is, with supra-contract norms. Relational contract theory postulates that where the ten common contract norms are inadequately served, exchange relations of whatever kind will fall apart.

Theory of core competencies is a popular theory in most public and private organisations. According to Prahalad and Hamel (1990), core competencies are the collective learning in the organisation, especially how to coordinate diverse production skills and integrate multiple streams of technologies. Since core competence is about harmonizing streams of technology, it is also about the organisation of work and the delivery of value.

Relational exchange theory is based on relational norms. Contracts are often extremely imperfect tools for controlling opportunism. While relational contracts may mitigate some opportunistic behavior, significant residual opportunism may remain. It is possible that transactors using relational contracts may incur significant ex-post bargaining costs as they periodically negotiate contract adjustments (Artz & Brush, 2000).

Stakeholder theory implies that the identification of stakeholders and their needs is important for decision-making in organisations. A stakeholder is any group or individual who can affect, or is affected by, the achievement of a corporation's purpose. Stakeholder theory is distinct because it addresses morals and values explicitly as a central feature of managing organisations. The ends of cooperative activity and the means of achieving these ends are critically examined in stakeholder theory in a way that they are not in many theories of strategic management (Phillips, Freeman, &

Wicks, 2003). According to Archmann and Kudlacek (2008), interoperability is not an end itself, but a tool to solve the problems of different stakeholders.

Theory of organisational boundaries claims that the resource-based view, transaction costs, and options perspectives each explain only a portion of managerial motivation for decisions on organisation boundaries. The rationale supporting the choices organisations make regarding member sourcing is multidimensional; firms are not only seeking potential sources of competitive advantage, but are also seeking to avoid opportunism and to preserve or create flexibility. There has been renewed debate on the determinants of firm boundaries and their implications for performance. According to Schilling and Steensma (2002), the widely accepted framework of transaction cost economics has come under scrutiny as a comprehensive theory for firm scale and scope. At the heart of this debate is whether the underlying mechanism determining firm boundaries is a fear of opportunism (as posited by transaction cost economics), a quest for sustainable advantage (as posed by resource-based view theorists and others), a desire for risk-reducing flexibility (as has recently gained increased attention in work on options), or a combination of factors. Although perspectives on firm boundaries such as transaction costs or the resource-based view are based on fundamentally different motivations for pursuing hierarchical control over market contracts, they rely on common resource or context attributes as antecedents.

Social exchange theory was initially developed to examine interpersonal exchanges that are not purely economic. Several sociologists are responsible for the early development of this theory. These theorists view people's social behavior in terms of exchanges of resources. The need for social exchange is created by the scarcity of resources, prompting actors to engage one another to obtain valuable inputs. Social exchange can be defined as voluntary actions of individuals that are motivated by return they are expected to bring and typically in fact bring from others. Social exchange can be viewed as an ongoing reciprocal process in which actions are contingent on rewarding reactions from others. There are important differences between social exchanges and economic exchanges. Social exchanges may or may not involve extrinsic benefits with objective economic value. In contrast to economic exchanges, the benefits from social exchanges often are not contracted explicitly, and it is voluntary to provide benefits. As a result, exchange partners are uncertain whether they will receive benefits. Thus social exchange theory focuses on the social relations among the actors that shape the exchange of resources and benefits. While its origins are at the individual level, social exchange theory has been extended to organisational and inter-organisational levels (Das & Teng, 2002a).

Institutional theory proposes that there is three institutional pillars that shape and constrain human behavior defined as the cognitive, normative and regulative structures and activities that provide stability and meanings to social behavior (Scott, 2001). The institutional pillars are universally defined and provide a lens for analyzing formal and informal inter-organisational differences and similarities which is prone to influence the ability to collaborate effectively and efficiently. Institutional theory has been applied in i.e. studies focusing on how institutional differences, either external to the company or internal, give rise to misunderstandings and institutional exceptions which in turn can lead to transaction costs.

Organisational climate (Taylor & Bowers, 1972) explicitly focuses on comparison between different social settings with the belief that a higher degree of inter-organisational similarities, mostly related to work practices, influence performance of inter-organisational cooperation. This contrasts somehow to the issue of organisation culture, which traditionally focuses on unique aspects of one social setting. Organisational climate research can be operationalized e.g. using “The Survey of Organisations” which supports more than 850 000 respondents (Bowers, 1988) which focus on the following domains;

organisation and work, communication flow, emphasis on human resource primacy, decision making practices and influence and control. The survey design presumes that certain social processes and relationships are common to all organisations and have a consistent correlation with performance and effectiveness. This drawback however might be less severe considering the focus on inter-operability in relatively similar social settings, compared to comparing organisation climate in organisations in different countries.

13.3.3 An example of empirical research

The Healthy Change Process Index (HCPI) as been developed in collaboration between NTNU, SINTEF and Arbeidsdirektoratet, and is a survey which aim to take the diagnosis of an organisations propensity to undergo a healthy change process based on several years of case studies on successful change process in Norwegian organisations. The healthier the change process is, the more likely it is to be effective. The Figure below show a guide for a healthy change processes using the HCPI application.

The first step is to probe the conditions for the change process, in which the HCPI will be used, incorporating the most crucial aspects of change processes. Statistical analyses of this data provide an indication of where resources for improvement should be aimed before either initiating a change process or continuing with current change process. It is advisable to take several measures during the change process as organisational change might take place and employees might find themselves in a different situation making them feel different about the structure and environment that they (now) function in.

HCPI measures healthy organisational change by the following success criteria:

- Awareness of norms
- Awareness of diversity
- Manager availability
- Role clarification
- Constructive conflicts

These are areas of major importance considering a healthy change process and reveals lacks and holes to be filled, or tells whether the organisation is ready for change or not, and also how they are doing

throughout the change process that usually lasts for some time depending on the nature of the change content. Also the HCPI measures *individual and collective resistance, involvement and information*. These data will reveal whether there at the given time exists a resistance towards the change among the employees, how involved the employees are in the change process due to both personal engagement and efforts made by the management to involve the employees in the process.

13.3.4 Summary and conclusions

E-government is a policy and managerial concept for which we have relatively little theoretical research. There is, however, a vast amount of empirical research available that focuses on the effects of ICT on the functioning of public administration in general (Bekkers & Homburg, 2007).
In our literature review (see Appendix 1 "Literature review e-Government interoperability") we found reports from several case studies. Examples of ICT applications studied are such as connecting citizens to pension reform in Sweden (Ranerup, 2006), Geneva's e-voting system (Chevallier, Warynski, & Sandoz, 2007), e-government impact on Shanghai firms' information process (Cui et al., 2006), and using geographic information systems locating crime hot spots for law enforcement (Gottschalk & Tolloczko, 2007). Typical topics covered are such as critical success factors and barriers to e-government, stakeholders involved in e-government projects, and conceptual frameworks.
In a research agenda for e-government integration and interoperability, Scholl and Klischewski (2007) suggest future research projects to study the foci and purposes, limitations and constraints, as well as processes and outcomes of integration and interoperation in electronic government.
For further research into organisational interoperability, the Semicolon project will look into barriers of e-government, maturity models of e-government interoperability, and we will try to develop theory-based benchmark variables to find different costs (and benefits) at different levels of organisational interoperability.

13.4 Information Governance

As defined by the UK NHS (National Health System) [1], Information Governance ensures necessary safeguards for, and appropriate use of, all information assets relevant for operational and financial valuation of an enterprise. This encompasses a number of issues tied to financial, cultural and legal aspects of organisational interoperability.

The current focus in the field of Information Governance seems divided between Data Governance, Information Security Governance, IT Governance and Internet Governance. In the following subsections we briefly describe each subfield. In the next chapter we outline a few examples of best practise activities.

13.4.1 Data Governance

Data Governance is a quality control discipline for assessing, managing, using, improving, monitoring, maintaining, and protecting organisational information [2] It is a system of decision rights and accountabilities for information-related processes, executed according to agreed-upon models which

describe who can take what actions with what information, and when, under what circumstances, using what methods [3]:

Data governance encompasses the people, processes, and information technology required to create a consistent and proper handling of an organisation's data across the business enterprise, including the goals of:

- Increasing consistency and confidence in decision making
- Decreasing the risk of regulatory fines
- Improving data security
- Maximizing the income generation potential of data
- Designating accountability for information quality

These goals are realized by the implementation of Data governance programs, or initiatives to improve data quality by assigning a team responsibility for data's accuracy, accessibility, consistency, and completeness, among other metrics. This team usually consists of executive leadership, project management, line-of-business managers, and data stewards. The team usually employs some form of methodology for tracking and improving enterprise data, such as Six Sigma, and tools for data mapping, profiling, cleansing, and monitoring data.

Data governance initiatives may be aimed at achieving a number of objectives including offering better visibility to internal and external customers (such as supply chain management), compliance with regulatory law, improving operations after rapid company growth or corporate mergers, or to aid the efficiency of enterprise knowledge workers by reducing confusion and error and increasing their scope of knowledge. Many data governance initiatives are also inspired by past attempts to fix information quality at the departmental level, leading to incongruent and redundant data quality processes. Most large companies have many applications and databases that can't easily share information. Therefore, knowledge workers within large organisations often don't have access to the information they need to best do their jobs. When they do have access to the data, the data quality may be poor. By setting up a data governance practice or Corporate Data Authority, these problems can be mitigated.

The structure of a data governance initiative will vary not only with the size of the organisation, but with the desired objectives or the 'focus areas' [3] of the effort.

13.4.2 Information Security Governance

Information security means protecting information and information systems from unauthorized access, use, disclosure, disruption, modification, or destruction [4]. The terms information security and computer security are frequently incorrectly used interchangeably. Information security is concerned with the confidentiality, integrity and availability of data regardless of the form the data may take: electronic, print, or other forms. Computer security can focus on ensuring the availability and correct operation of a computer system without concern for the information stored or processed by the computer.

Confidentiality is the property of preventing disclosure of information to unauthorized individuals or systems. For example, a credit card transaction on the Internet requires the credit card number to be transmitted from the buyer to the merchant and from the merchant to a transaction processing network. The system attempts to enforce confidentiality by encrypting the card number during transmission, by

limiting the places where it might appear (in databases, log files, backups, printed receipts, and so on), and by restricting access to the places where it is stored. If an unauthorized party obtains the card number in any way, a breach of confidentiality has occurred.

In information security, integrity means that data cannot be modified without authorization. Integrity is violated when an employee (accidentally or with malicious intent) deletes important data files, when a computer virus infects a computer, when an employee is able to modify his own salary in a payroll database, when an unauthorized user vandalizes a web site, when someone is able to cast a very large number of votes in an online poll, and so on.

For an information system to serve its purpose, availability is required such that information can be used when needed. This means that the computing systems used to store and process the information, the security controls used to protect it, and the communication channels used to access it must be functioning correctly. High availability systems aim to remain available at all times, preventing service disruptions due to power outages, hardware failures, and system upgrades. Ensuring availability also involves preventing denial-of-service attacks.

In 2002, Donn Parker proposed an alternative model for the classic CIA triad (confidentiality, integrity, availability) that he called the six atomic elements of information. The elements are confidentiality, possession, integrity, authenticity, availability, and utility. The merits of the Parkerian hexad are a subject of debate amongst security professionals.

An important aspect of information security and risk management is recognizing the value of information and defining appropriate procedures and protection requirements for the information. Not all information is equal and so not all information requires the same degree of protection. This requires information to be assigned a security classification.

The first step in information classification is to identify a member of senior management as the owner of the particular information to be classified. Next, develop a classification policy. The policy should describe the different classification labels, define the criteria for information to be assigned a particular label, and list the required security controls for each classification.

Some factors that influence which classification information should be assigned include how much value that information has to the organisation, how old the information is and whether or not the information has become obsolete. Laws and other regulatory requirements are also important considerations when classifying information. Common information security classification labels used by the business sector are: *public, sensitive, private, confidential.* Common information security classification labels used by government are: *Unclassified, Sensitive But Unclassified, Restricted, Confidential, Secret, Top Secret* and their non-English equivalents.

All employees in the organisation, as well as business partners, must be trained on the classification schema and understand the required security controls and handling procedures for each classification. The classification a particular information asset has been assigned should be reviewed periodically to ensure the classification is still appropriate for the information and to ensure the security controls required by the classification are in place.

Access control: Protected information must be restricted to people who are authorized to access the information. Access to the computer programs, and in many cases the computers that process the information, must also be authorized. This requires that mechanisms be in place to control the access to protected information. The sophistication of the access control mechanisms should be in parity with the value of the information being protected - the more sensitive or valuable the information the stronger the control mechanisms need to be. The foundation on which access control mechanisms are built start with identification and authentication.

Identification is an assertion of who someone is or what something is. If a person makes the statement *"Hello, my name is John Doe."* they are making a claim of who they are. However, their claim may or may not be true. Before John Doe can be granted access to protected information it will be necessary to verify that the person claiming to be John Doe really is John Doe.

Authentication is the act of verifying a claim of identity. When John Doe goes into a bank to make a withdrawal, he tells the bank teller he is John Doe (a claim of identity). The bank teller asks to see a photo ID, so he hands the teller his driver's license. The bank teller checks the license to make sure it has John Doe printed on it and compares the photograph on the license against the person claiming to be John Doe. If the photo and name match the person, then the teller has authenticated that John Doe is who he claimed to be.

There are three different types of information that can be used for authentication: *something you know, something you have, or something you are***.** Examples of *something you know* include such things as a PIN, a password, or your mother's maiden name. Examples of *something you have* include a driver's license or a magnetic swipe card. *Something you are* refers to biometrics. Examples of biometrics include palm prints, finger prints, voice prints and retina (eye) scans. Strong authentication requires providing information from two of the three different types of authentication information. For example, something you know plus something you have. This is called two factor authentication.

13.4.3 Information Technology Governance

Information Technology Governance, IT Governance or ICT (Information & Communications Technology) Governance, is a subset discipline of Corporate Governance focused on information technology (IT) systems and their performance and risk management. The rising interest in IT governance is partly due to compliance initiatives, for instance Sarbanes-Oxley in the USA and Basel II in Europe, as well as the acknowledgment that IT projects can easily get out of control and profoundly affect the performance of an organisation.

A characteristic theme of IT governance discussions is that the IT capability can no longer be a black box. The traditional involvement of board-level executives in IT issues was to defer all key decisions to the company's IT professionals. IT governance implies a system in which all stakeholders, including the board, internal customers, and in particular departments such as finance, have the necessary input into the decision making process. This prevents IT from independently making and later being held solely responsible for poor decisions. It also prevents critical users from later complaining that the system does not behave or perform as expected, as explained in the Harvard Business Review article by R. Nolan:

A board needs to understand the overall architecture of its company's IT applications portfolio … The board must ensure that management knows what information resources are out there, what condition they are in, and what role they play in generating revenue…[5].

There are narrower and broader definitions of IT governance. Weill and Ross focus on "Specifying the decision rights and accountability framework to encourage desirable behaviour in the use of IT." [6]

In contrast, the IT Governance Institute expands the definition to include foundational mechanisms: "*… the leadership and organisational structures and processes that ensure that the organisation's IT sustains and extends the organisation's strategies and objectives.*" [7]

13.4.4 Internet Governance

The definition of Internet governance has been contested by differing groups across political and ideological lines. One of the key debates centers on the authority and participation of certain actors, such as national governments and corporate entities, to play a role in the Internet's governance.

A Working Group established after a United Nations-initiated World Summit on the Information Society (WSIS) proposed the following definition of Internet governance as part of its June 2005 report:

Internet governance is the development and application by Governments, the private sector and civil society, in their respective roles, of shared principles, norms, rules, decision-making procedures, and programmes that shape the evolution and use of the Internet [8]

Law professor Yochai Benkler developed a framework for conceptualizing the idea of Internet governance through the idea of three "layers" of governance: the "physical infrastructure" layer through which information travels; the "code" or "logical" layer that controls the infrastructure; and the "content" layer, which contains the information that runs through the network [9].

14 BEST PRACTICE

14.1 Best practise on Information Governance

In this chapter we and outline a few examples of best practise activities. Needless to say, these are only a small sample of the many activities and initiatives that take place worldwide on different aspects of Information Governance, and many more references to ongoing work can be found from information published and updated on (an almost daily basis on) the Internet.

14.1.1 Implementation of Data Governance

Leaders of successful data governance programs declared in December 2006 at the Data Governance Conference in Orlando, Fl, that data governance is between 80 and 95 percent communication. [10] That stated, it is given that many of the objectives of a Data Governance program must be accomplished with appropriate tools. Many vendors are now positioning their products as Data Governance tools; due to the different focus areas of various data governance initiatives, any given tool may or may not be appropriate, in addition, many tools that are not marketed as governance tools address governance needs [11].

The IBM Data Governance Council [12] is an organisation formed by IBM consisting of companies, institutions and technology solution providers with the stated objective to build consistency and quality control in governance, which will help companies better protect critical data."

The Data Governance and Stewardship Community of Practice (DGS-COP) [13] is a vendor-neutral organisation open to practitioners, stakeholders and academics, as well as vendors and consultants. The DGS-COP offers a large collection of data governance artifacts to members including case studies, metrics, dashboards, and maturity models as well as on-line events.

Data Governance Conferences [14] Two major conferences are held annually, the Data Governance Conference, held in 2008 in San Francisco, CA, USA, and the Data Governance Conference Europe, held in 2008 in London, England.

Implementation of a Data Governance initiative may vary in scope as well as origin. Sometimes, an executive mandate will arise to initiate an enterprise wide effort, sometimes the mandate will be to create a pilot project or projects, limited in scope and objectives, aimed at either resolving existing issues or demonstrating value. Sometimes an initiative will originate lower down in the organisation's hierarchy, and will be deployed in a limited scope to demonstrate value to potential sponsors higher up in the organisation.

14.1.2 Implementation of Information Security Governance

ISO/IEC 27002 *Information technology - Security techniques - Code of practice for information security management* is an information security standard, and part of a growing family of ISO/IEC ISMS standards (the ISO/IEC 27000 series). It was published by the International Organisation for Standardization (ISO) and the International Electrotechnical Commission (IEC) as ISO/IEC 17799:2005 and subsequently renumbered ISO/IEC 27002:2005 in July 2007, bringing it into line with the other ISO/IEC 27000-series standards. The current standard is a revision of the version first published by ISO/IEC in 2000, which was a word-for-word copy of the British Standard (BS) 7799-1:1999.

ISO/IEC 27002 provides best practice recommendations on information security management for use by those who are responsible for initiating, implementing or maintaining Information Security Management Systems (ISMS). Information security is defined within the standard in the context of the C-I-A triad:

the preservation of confidentiality (ensuring that information is accessible only to those authorized to have access), integrity (safeguarding the accuracy and completeness of information and processing methods) and availability (ensuring that authorized users have access to information and associated assets when required).

ISO 17799:2005 standard is the most recently published revision of ISO's global security framework. This version significantly improves the already well-respected and comprehensive "Code of Practice for Information Security Management." It provides principles and guidelines for initiating, implementing, maintaining, and improving information security management throughout the enterprise. This includes best practices, control objectives and controls for a range of IT functions

related to protecting information.

The ISO/IEC 27002:2005 Code of practice for information security management recommends the following be examined during a risk assessment:

- security policy,
- organisation of information security,
- asset management, human resources security,
- physical and environmental security,
- communications and operations management,
- access control,
- information systems acquisition,
- development and maintenance,
- information security incident management,
- business continuity management, and
- regulatory compliance.

With respect to identification and authentication for computer systems in use today, the Username is the most common form of identification and the Password is the most common form of authentication. Usernames and passwords have served their purpose but in our modern world they are no longer adequate. Usernames and passwords are slowly being replaced with more sophisticated authentication mechanisms.

After a person, program or computer has successfully been identified and authenticated then it must be determined what informational resources they are permitted to access and what actions they will be allowed to perform (run, view, create, delete, or change). Such *authorization* to access information and other computing services begins with administrative policies and procedures. The policies prescribe what information and computing services can be accessed, by whom, and under what conditions. The access control mechanisms are then configured to enforce these policies. Different computing systems are equipped with different kinds of access control mechanisms; some may offer a choice of different access control mechanisms. The access control mechanism a system offers will be based upon one of three approaches to access control or it may be derived from a combination of the three approaches.

The non-discretionary approach consolidates all access control under a centralized administration. The access to information and other resources is usually based on the individuals function (role) in the organisation or the tasks the individual must perform. The discretionary approach gives the creator or owner of the information resource the ability to control access to those resources. In the Mandatory access control approach, access is granted or denied bases upon the security classification assigned to the information resource.

Examples of common access control mechanisms in use today include Role-based access control available in many advanced Database Management Systems, simple file permissions provided in the UNIX and Windows operating systems, Group Policy Objects provided in Windows network systems, Kerberos, RADIUS, TACACS, and the simple access lists used in many firewalls and routers.

To be effective, policies and other security controls must be enforceable and upheld. Effective policies ensure that people are held **accountable** for their actions. All failed and successful authentication attempts must be logged, and all access to information must leave some type of audit trail.

In computing, e-Business and information security it is necessary to ensure that the data, transactions, communications or documents (electronic or physical) are genuine (i.e. they have not been forged or fabricated.). It is also important for authenticity to validate that both parties involved are who they claim they are.

In law, non-repudiation implies one's intention to fulfill their obligations to a contract. It also implies that one party of a transaction can not deny having received a transaction nor can the other party deny having sent a transaction.

Electronic commerce uses technology such as digital signatures and encryption to establish authenticity and non-repudiation. Digital signatures are also used in healthcare sector for communication containing sensitive health information.

14.1.3 Implementation of Information Technology Governance

After the widely reported collapse of Enron in 2000, and the alleged problems within Arthur Andersen and WorldCom, the duties and responsibilities of the boards of directors for public and privately held corporations were questioned. As a response to this, and to attempt to prevent similar problems from happening again, the US Sarbanes-Oxley Act (SOX) was written to stress the importance of business control and auditing. Sarbanes-Oxley and Basel-II in Europe have been catalysts for the development of the discipline of information technology governance since the early 2000s. However, the concerns of Sarbanes Oxley (in particular Section 404) have less to do with IT decision rights as discussed by Weill and Ross [6] and more to do with operational control processes such as Change management.

What IT controls are most important for SOX compliance? A growing number of corporate IT organisations are finding at least some of the answers in recent iterations of two venerable standards frameworks: COBIT and ISO 17799:2005 (renamed to ISO 27005, as described in the previous section on Information Security implementation).

Control Objectives for Information and related Technologies (COBIT) is an open standard published by the IT Governance Institute and the Information Systems Audit and Control Association (ISACA).[15] A new version recently published, COBIT 4.0, emphasizes regulatory compliance as it relates to IT governance. ISACA, and describes COBIT as an IT governance framework with a supporting toolset that allows managers to bridge the gap between control requirements, technical issues and business risks.

COBIT provides a best practice framework for how to control, manage and measure 34 key IT practices. This framework includes high-level and detailed control objectives for each process, management guidelines (including process inputs and outputs, roles and responsibilities, and metrics), and process maturity models. A core emphasis of COBIT is aligning IT operations with strategic enterprise objectives and priorities to improve IT value delivery, resource management, business performance, efficiency and risk management.

The ISO 17799:2005 standard includes extensions that strengthen controls designed to protect the integrity of information from asset management and access control, to human resources security, security incident management and business continuity management. An important new requirement is

an increased emphasis not only on the need to have good security controls, but also on the capability to validate the integrity of regulated information. It mandates validation through systematic auditing and monitoring of activity to prevent unauthorized access to sensitive corporate and customer information. Just as ISO 9000/9001 is used universally as a measure of production quality, ISO 17799:2005 is poised to play a similar role in the area of information integrity assurance.

Both COBIT and ISO 17799/2005 provide guidelines that are useful in helping companies determine how to think about the root requirements of compliance regulations and managing data risks. Developed specifically for IT organisations, these frameworks provide specific practices and guidelines for instituting controls aimed at ensuring the integrity of information assets.

Following Corporate Collapses in Australia around the same time, working groups were established to develop standards for Corporate Governance. A series of Australian Standards for Corporate Governance were published in 2003, these were:
Good Governance Principles (AS8000)
Fraud and Corruption Control (AS8001)
Organisational Codes of Conduct (AS8002)
Corporate Social Responsibility (AS8003)
Whistle Blower protection programs (AS8004)

AS8015 Corporate Governance of ICT, the Australian Standard for Corporate Governance of ICT, was published in January 2005. It was fast-track adopted as ISO/IEC 38500 in May 2008. AS8015 defines Corporate Governance of ICT as "The system by which the current and future use of ICT is directed and controlled. It involves evaluating and directing the plans for the use of ICT to support the organisation and monitoring this use to achieve plans. It includes the strategy and policies for using ICT within an organisation."

ISO/IEC 29382, Corporate Governance of Information and Communication Technology, was first published early in 2007 as a fast track candidate from the existing Australian standard AS8015. It was officially re-named ISO/IEC 38500 in April 2008. As is usual with international standards, it is intended to provide guiding principles to any organisation, regardless of size or sector.

ISO 38500 is now the international standard for the corporate governance of information technology has now been published. The original draft number for the standard of ISO 29382 has been discarded, and the official number of the new standard is ISO/IEC 38500. It draws upon a number of sources, chief of which is AS 8015:2005, which defines six principles (establish responsibilities, plan to best support the organisation, acquire validly, ensure performance when required, ensure conformance with rules, ensure respect for human factors).

ISO/IEC 38500:2008, corporate governance of information technology, is applicable to organisations of all sizes, including public and private companies, government entities, and not-for-profit organisations. This standard provides a framework for effective governance of IT to assist those at the highest level of organisations to understand and fulfil their legal, regulatory and ethical obligations in respect of their organisations' use of IT. The framework comprises definitions, principles and a model.

It sets out six principles for good corporate governance of IT that express preferred behaviour to guide decision making:
(1) responsibility, (2) strategy, (3) acquisition, (4) performance, (5) conformance and (6) human behaviour.

The purpose of the standard is to promote effective, efficient and acceptable use of IT in all organisations by: assuring stakeholders that, if the standard is followed, they can have confidence in the organisation's corporate governance of IT informing and guiding directors in governing the use of IT in their organisation, and providing a basis for objective evaluation of the corporate governance of IT.

Practical application of these standards and principles typically takes place in IT Service Management (ITSM) programs, which is a discipline for managing information technology (IT) systems, philosophically centered on the *customer's perspective of IT's contribution to the business.* ITSM stands in deliberate contrast to technology-centered approaches to IT management and business interaction. ITSM is often equated with the Information Technology Infrastructure Library, (ITIL), an official publication of the Office of Government Commerce [16] in the United Kingdom. However, while a version of ITSM is a component of ITIL, ITIL also covers a number of related but distinct disciplines and the two are not synonymous.

ITIL is a set of concepts and policies for managing information technology (IT) infrastructure, development and operations. ITIL gives a detailed description of a number of important IT practices with comprehensive check lists, tasks and procedures that can be tailored to any IT organisation. The "Service Management" section of ITIL version 2 was made up of eleven different disciplines, split into two sections, Service Support and Service Delivery. This use of the term "Service Management" is how many in the world interpret ITSM, but again, there are other frameworks, and conversely, the entire ITIL library might be seen as IT Service Management in a larger sense. The new ITIL v3 rewrite has not similarly designated a subset as "Service Management."

14.1.4 Implementation of Internet Governance

To understand how the Internet is run today, it is necessary to know some of the key milestones of Internet governance.

The original ARPANET, one of the components which eventually evolved into the Internet, connected four Universities: University of California Los Angeles, University of California Santa Barbara , Stanford Research Institute and Utah University. The IMPs, interface minicomputers, were built in 1969 by Bolt, Beranek and Newman under a proposal by the US Department of Defense Advanced Research Projects Agency. By 1973 the ARPANET connected many more systems and included satellite links to Hawaii and Scandinavia, and a further link from Norway to London. It continued to grow in size, becoming more a utility than a research project. For this reason in 1975 it was transferred to the US Defense Communications Agency. During the development of ARPANET, a numbered series of Request for Comments (RFCs) memos documented technical decisions and methods of working as they evolved. The standards of today's Internet are still documented by RFCs, produced through the very process which evolved on ARPANET. The Internet protocol suite, developed

between 1973 and 1977 with funding from ARPA, was intended to hide the differences between different underlying networks and allow many different applications to be used over the same network.

In 1979 the Internet Configuration Control Board was founded by DARPA to oversee the network's development. In 1984 it was renamed the Internet Advisory Board (IAB), and in 1986 it became the Internet Activities Board. RFC 801 describes how the US Department of Defense organized the replacement of ARPANET's Network Control Program by the new Internet Protocol in January 1983. In the same year, the military systems were removed to a distinct MILNET, and the Domain Name System was invented to manage the names and addresses of computers on the "ARPA Internet". The familiar top-level domains .gov, .mil, .edu, .org, .net, .com, and .int, and the two-letter country code top-level domains were deployed in 1984. Between 1984 and 1986 the US National Science Foundation created the NSFNET backbone, using TCP/IP, to connect their supercomputing centers. The combined network became widely known as the Internet.

Outside of the USA the dominant technology was X.25. The International Packet Switched Service, created in 1978, used X.25 and extended to Europe, Australia, Hong Kong, Canada, and the USA. It allowed individual users and companies to connect to a variety of mainframe systems, including Compuserve. Between 1979 and 1984, an approach known as Unix to Unix Copy Program grew to connect 940 hosts, using methods like X.25 links, ARPANET connections, and leased lines. Usenet News, a distributed discussion system, was a major use of UUCP. The Internet Engineering Task Force (IETF) was formed in 1986 by the US Government to develop and promote Internet standards. It initially consisted of researchers, but by the end of the year participation was open to anyone, and its business was largely carried on by email. By the end of 1989 Australia, Germany, Israel, Italy, Japan, Mexico, the Netherlands, New Zealand, and the United Kingdom had connected to the Internet, which now contained over 160,000 hosts.

In 1990, ARPANET formally shut down, and in 1991 the NSF dropped its restrictions on commercial use of its part of the Internet. Commercial network providers began to interconnect, extending the Internet. In 1992 the Internet Society (ISOC) was founded, with a mission to *"assure the open development, evolution and use of the Internet for the benefit of all people throughout the world"*. Its members include individuals (anyone may join) as well as corporations, organisations, governments, and universities. The IAB was renamed the Internet *Architecture* Board, and became part of ISOC. The Internet Engineering Task Force also came under the ISOC umbrella. The IETF is currently overseen by the Internet Engineering Steering Group (IESG), and longer term research is carried on by the Internet Research Task Force and overseen by the Internet Research Steering Group.

Allocation of IP addresses was delegated to four Regional Internet Registries (RIRs):
American Registry for Internet Numbers (ARIN) for North America
Réseaux IP Européens - Network Coordination Centre (RIPE NCC) for Europe, the Middle East, and Central Asia Asia-Pacific Network Information Centre (APNIC) for Asia and the Pacific region Latin American and Caribbean Internet Addresses Registry (LACNIC) for Latin America and the Caribbean region In 1998, the IANA function was taken over by the Internet Corporation for Assigned Names and Numbers (ICANN), a newly created Californian non-profit corporation, set up in September 1998 by the US Government and awarded a contract by the US Department of Commerce. Initially two board members were elected by the Internet community at large, though this was changed by the rest of the board in 2002 in a thinly attended public meeting in Accra, in Ghana. In 2004 a new RIR,

AfriNIC, was created to manage allocations for Africa. In 2002, a restructuring of the Internet Society gave more control to its corporate members.

At the first World Summit on the Information Society (WSIS) in Geneva 2003 the topic of Internet governance was put on the table. Since no general agreement existed even on the definition of what comprised Internet governance, United Nations Secretary General Kofi Annan set up a Working Group on Internet Governance (WGIG) to clarify the issues and report before the second part of the World Summit on the Information Society in Tunis 2005. After much controversial debate, participants agreed on a compromise to allow for wider international debate on the policy principles. They agreed to establish an Internet Governance Forum, to be convened by United Nations Secretary General before the end of the second quarter of the year 2006. The Greek government volunteered to host the first such meeting.

Today almost all Internet infrastructure is provided and owned by the private sector. Traffic is exchanged between these networks, at major interconnect points, in accordance with established Internet standards and commercial agreements.

15 REFERENCES

[Thompson] Thompson, James D. [1967] (2003). Organizations in Action: Social Science Bases of Administrative Theory,

[Galbraith] Jay Galbraith, Designing Complex Organizations – An Information Processing View, 1974, Interfaces, United States

[Scott] W. Richard Scott, Organizations: Rational, Natural, and Open Systems, 4th Ed. (Upper Saddle River, NJ: Prentice Hall, 1998),

[Christiansen] Christiansen, T.R. "Modeling efficiency and effectiveness of coordination in engineering design teams" PhD Thesis Dissertation, Department of Civil Engineering, Stanford University, California, 1993 - Published as DNV Research report 93-2063

[Smith-Meyer, 2005] H.Smith-Meyer, F.Lillehagen, H.D.Jørgensen, D.Karlsen, E.Dehli, S.Murberg, H.G.Solheim : Using Models and Model-based Transformation to Generate Operational Context-sensitive Workplaces. *Proceedings Collaborative Engineering 2005.*

[1] Afuah, A., & Tucci, C. L. (2003). *Internet Business Models and Strategies* (2nd ed.). New York, NY: McGraw-Hill.

[2] Anderson, S. W., Glenn, D., & Sedatole, K. L. (2000). Sourcing parts of complex products: evidence on transactions costs, high-powered incentives and ex-postopportunism. *Accounting, Organisations and Society, 25*(5), 723-749.

[3] Archmann, S., & Kudlacek, I. (2008). Interoperability and the exchange of good practice cases. *European Journal of ePractice, 2*(February), 3-12.

[4] Artz, K. W., & Brush, T. H. (2000). Asset specificity, uncertainty and relational norms: an examination of coordination costs in collaborative strategic alliances. *Journal of Economic Behavior & Organisation, 41*(4), 337-362.

[5] Bekkers, V. (2007). The government of back-office integration. Organizing co-operation between information domains. *Public Management Review, 9*(3), 377-300.

[6] Bekkers, V., & Homburg, V. (2007). The Myths of E-Government: Looking beyond the Assumptions of a New and Better Government. *Information Society, 23*(5), 373-382.

[7] Bowers, D. (1988). *The survey of organisations*. Ann Arbor, MI: Rensis Likert and ass.

[8] Chen, Z., Gangopadhyay, A., Holden, S. H., Karabatis, G., & McGuire, P. (2007). Semantic integration of government data for water quality management. *Government Information Quarterly, 24*(4), 716-735.

[10] Chevallier, M., Warynski, M., & Sandoz, A. (2007). Success factors of Geneva's e-Voting system. *European Journal of ePractice, 1*(November), 47-59.

[11] Clarkson, G., Jacobsen, T. E., & Batcheller, A. L. (2007). Information asymmetry and information sharing. *Government Information Quarterly, 24*(4), 827-839.

[12] Cui, L., Zhang, C., Zhang, C., & Huang, L. (2006). Exploring E-Government Impact on Shanghai Firms' Information Process. *Electronic Markets, 16*(4), 312-328.

[13] Das, T. K., & Teng, B.-S. (2002a). Alliance Constellations: A Social Exchange Perspective. *Academy of Management Review, 27*(3), 445-456.

[14] Das, T. K., & Teng, B.-S. (2002b). The dynamics of alliance conditions in the alliance development process. *Journal of Management Studies, 39*(5), 725-746.

[15] Das, T. K., & Teng, B.-S. (2003). Partner analysis and alliance performance. *Scandinavian Journal of Management, 19*(3), 279-308.

[16] Eckman, B. A., Bennet, C. A., Kaufman, J. H., & Tenner, J. W. (2007). Varieties of interoperability in the transformation of the health-care information infrastructure. *IBM Systems Journal, 46*(1), 19-41.

[17] Eisenhardt, K. M. (1985). Control: organisational and economic approaches. *Management Science, 31*(2), 134-149.

[18] Enyon, R., & Margetts, H. (2007). Organisational Solutions for Overcoming Barriers to eGovernment. *European Journal of ePractice, 1*(November), 73-85.

[19] Esteves, J., & Joseph, R. C. (2008). A comprehensive framework for the assessment of eGovernment projects. *Government Information Quarterly, 25*(1), 404-437.

[20] Fedorowicz, J., Gogan, J. L., & Williams, C. B. (2007). A collaborative network for first responders: Lessons from CapWIN case. *Government Information Quarterly, 24*(4), 785-807.

[21] Fléri, L. F., Moutet, G., & Cimander, R. (2007). Regional Shared eGovernment in the Region of Burgundy: the case of eProcurement. *European Journal of ePractice, 1*(November), 60-72.

[22] Gogan, J. L., Williams, C. B., & Fedorowicz, J. (2007). RFID and interorganisational collaboration: political and administrative challenges. *Electronic Government, an International Journal, 4*(4), 423-435.

[23] Gottschalk, P., & Solli-Sæther, H. (2008). Stages of e-government interoperability. *Electronic Government, an International Journal, 5*(3), 310-320.

[24] Gottschalk, P., & Tolloczko, P. (2007). Maturity model for mapping crime in law enforcement. *Electronic Government, an International Journal, 4*(1), 59-67.

[25] Gouscos, D., Kalikakis, M., Legal, M., & Papadopoulou, S. (2007). A general model of performance and quality for one-stop e-government service offerings. *Government Information Quarterly, 24*(4), 860-885.

[26] Grant, G., McKnight, S., Uruthirapathy, A., & Brown, A. (2007). Designing governance for shared services organisations in the public service. *Government Information Quarterly, 24*(3), 533-538.

[27] Grefen, P., Ludwig, H., & Angelov, S. (2003). A three-level framework for process and data management of complex e-services. *International Journal of Cooperative Information Systems, 12*(4), 487-531.

[28] Greiner, U., Legner, C., Lippe, S., & Wende, K. (2007). Business Interoperability Profiles. In R. Goncales, J. P. Müller, K. Mertins & M. Zelm (Eds.), *Enterprise Interoperability II* (pp. 865-877). London, UK: Springer Verlag.

[29] Jensen, M. C., & Meckling, W. H. (1976). Theory of the firm: Managerial behavior, agency costs and ownership structures. *Journal of Financial Economics, 3*(4), 305-360.

[30] Klischewski, R., & Scholl, H. J. (2008). Information quality as capstone in negotiating e-government integration, interoperation and information sharing. *Electronic Government, an International Journal, 5*(2), 203-225.

[31] Legner, C., & Lebreton, B. (2007). Business Interoperability Research: Present Achievements and Upcoming Challenges. *Electronic Markets, 17*(3), 176-186.

[32] Luna-Reyes, L. F., Gil-Garcia, J. R., & Cruz, C. B. (2007). Collaborative digital government in Mexico: Some lessons from federal web-based interorganisational information integration initiatives. *Government Information Quarterly, 24*(4), 808-826.

[33] Luo, Y. (2002). Contract, cooperation, and performance in international joint ventures. *Strategic Management Journal, 23*(10), 903-919.

[34] Macneil, I. R. (2000). Relational contract theory: challenges and queries. *Northwestern University Law Review, 94*(3), 877-907.

[35] Otjacques, B., Hitzelberger, P., & Feltz, F. (2007). Interoperability of E-Government Information Systems: Issues of identification and Data Sharing. *Journal of Management Information Systems, 23*(4), 29-51.

[36] Papazoglou, T. A., & Ribbers, P. M. A. (2006). *E-Business: Organisational and Technical Foundations*. West Sussex, UK: John Wiley & Sons.

[37] Pardo, T. A., & Tayi, G. K. (2007). Interorganisational information integration: A key enabler for digital government. *Government Information Quarterly, 24*(4), 691-715.

[38] Pedersen, M. K., Fountain, J., & Loukis, E. (2006). Preface to the Focus Theme Section: 'Electronic Markets and e-Government'. *Electronic Markets, The International Journal, 16*(4), 263-273.

[39] Phillips, R., Freeman, R. E., & Wicks, A. C. (2003). What stakeholder theory is not. *Business Ethics Quarterly, 13*(4), 479-502.

[40] Prahalad, C. K., & Hamel, G. (1990). The Core Competence of the Corporation. *Harvard Business Review, 68*(3), 79-91.

[41] Ranerup, A. (2006). Electronic Markets Connecting Citizens to Pension Reform. *Electronic Markets, The International Journal, 16*(4), 282-291.

[42] Ranerup, A. (2007). Electronic Government as a combination of human and technological agency: Testing the principle os symmetry. *Information Polity: The International Journal of Gvoernment & Democracy in the Information Age, 12*(3), 153-167.

[43] Reddick, C. G. (2008). Collaborative management and e-government: a survey of state government CIOs. *Electronic Government, an International Journal, 5*(2), 146-161.

[44] Salleh, N. A. M., Rohde, F., & Green, P. (2006). The Effect of Enacted Capabilities on adoption of a Government Electronic Procurement System by Malaysian SMEs. *Electronic Markets, The International Journal, 16*(4), 292-311.

[45] Santos, E. M. D., & Reinhard, N. (2007). Setting interoperability standards for e-government: an exploratory case study. *Electronic Government, an International Journal, 4*(4), 889-920.

[46] Schilling, M. A., & Steensma, H. K. (2002). Disentangling the Theories of firm Boundaries: A Path Model and Empirical Test. *Organisation Science, 13*(4), 387-401.

[47] Scholl, H. J., & Klischewski, R. (2007). E-Government Integration and Interoperability: Framing the Research Agenda. *International Journal of Public Administration, 30*(8), 889-920.

[48] Schooley, B. L., & Horan, T. A. (2007). Towards end-to-end government performance management: Case study of interorganisational information integration in emergency medical services (EMS). *Government Information Quarterly, 24*(4), 755-784.

[49] Scott, W. R. (2001). *Institutions and Organisations* (2nd ed.). California: Sage.

[50] Taylor, J., & Bowers, D. (1972). *The Survey of Organisations*. Ann Arbor (MI): University of Michigan Press.

[51] United Nations. (2008). *UN e-Government Survey. From e-Government to Connected Governance* (No. ST/ESA/PAD/SER.E/112). New York: Department of Economics and Social Affairs, Divison for Public Administration and Development Management.

[52] Vragov, R., & Kumar, N. (2006). Electronic Markets for the Allocation, Financing and Distribution of Public Goods. *Electronic Markets, The International Journal, 16*(4), 274-281.

[53]Wang, H., Song, Y., Hamilton, A., & Curwell, S. (2007). Urban information integration for advanced e-Planning in Europe. *Government Information Quarterly, 24*(4), 736-754.

[54] Williamson, O. E. (1979). Transaction-Cost Economics: The Governance of Contractual Relations. *The Journal of Law and Economics, 22*(2), 233-261.

[55] http://www.connectingforhealth.nhs.uk/systemsandservices/infogov

[56] http://datagovernance.com/dgi_framework.pdf

[57] http://datagovernance.com/fc_focus_areas_for_data_governance.html

[58] http://www.law.cornell.edu/uscode/html/uscode44/usc_sec_44_00003542----000-.html

[59] Nolan, R. and F. W. McFarlan (2005). "Information Technology and the Board of Directors." *Harvard Business Review* (October 2005).

[60] Weill, P. & Ross, J. W., 2004, *IT Governance: How Top Performers Manage IT Decision Rights for Superior Results*, Harvard Business School Press, Boston.

[61] IT Governance Institute 2003, "Board Briefing on IT Governance, 2nd Edition". from http://www.isaca.org/Content/ContentGroups/ITGI3/Resources1/Board_Briefing_on_IT_Governance/26904_Board_Briefing_final.pdf

[62] WGIG (2005), p.4. Available at: http://www.wgig.org/docs/WGIGREPORT.pdf

[63] Yochai Benkler, *From Consumers to Users: Shifting the Deeper Structures of Regulation Towards Sustainable Commons and User Access*, 52 Fed. Comm. L.J. 561, (2000).

[64] Hopwood, Peter (2008-06). "Data Governance: One Size Does Not Fit All". DM Review Magazine. Archived from the original on 2008-10-02. Retrieved on 2008-10-02. "At the inaugural Data Governance Conference in Orlando, Florida, in December 2006, leaders of successful data governance programs declared that in their experience, data governance is between 80 and 95 percent communication. Clearly, data governance is not a typical IT project."

[65] "DataGovernanceSoftware.com". http://www.webcitation.org/5bGI3dfHV The Data Governance Institute. Archived from the original on 2008-10-02. Retrieved on 2008-10-02.

[66] IBM Data Governance http://www-01.ibm.com/software/tivoli/governance/servicemanagement/data-governance.html

[67] The Data Governance & Stewardship Community of Practice http://www.datastewardship.com/

[68] Data Governance 2008 Conference http://www.debtechint.com/dg2008/

[69]http://www.isaca.org/Content/ContentGroups/ITGI3/Resources1/Board_Briefing_on_IT_Governance/26904_Board_Briefing_final.pdf

[70] Office of Government Commerce (2001). *Service Delivery: Capacity Management, Availability Management, Service Level Management, IT Service Continuity, Financial Management for IT Services and Customer Relationship Management.* OGC, ITIL© Managing IT Services (IT Infrastructure Library). London, The Stationery Office. ISBN 0-11-330017-4

[71] Manuela Finetti (2003), *Interoperability matters*, Information technology

[72] IDABC (2005), EIF

[73] IDABC (2008), EIF, draft 2

[74] Victor Bekkers (2007), *The governance of back office integration*

[75] Eynon-Margetts (2007), Organisational solution for overcoming barriers to eGovernment, European Journalof ePractice

[76] Ministry of Finance (2005), *Veileder i samfunnsøkonomiske analyser*.

[77] Government Agency for Financial Management (2007), *Veileder i samfunnsøkonomisk analyse og gevinstrealisering av IKT-prosjekter.*

[78] Schartum (2008), *IT-støtte for arbeid med lovsaker*.

Additional references -

[78] Carr, N. G. (2004). *Does IT matter? : information technology and the corrosion of competitive advantage.* Boston, Harvard Business School Press. ISBN 1-59139-444-9

[79] Remenyi, D., A. H. Money, et al. (2000). *The effective measurement and management of IT costs and benefits.* Computer weekly professional series. Oxford ; Boston, Butterworth-Heinemann. ISBN 0-7506-4420-6

16 APPENDICES

16.1 Appendix 1: Literature review e-Government interoperability

Title/ authors	*Research method*	*Scope of the research*	*Key conclusions/ findings*
A three-level framework for process and data management of complex e-services (Grefen, Ludwig, & Angelov, 2003)	Conceptual	Development of a three-level process and data specification framework. The three levels are: internal, conceptual, and external level.	The three-level approach to business process specification provides a clear separation of concerns in business process design, thereby increasing quality, flexibility and reusability of process specifications in cross-organisational settings. Separation of concerns is becoming increasingly important, as the complexity of automated cross-organisational processes grows through the advent of digital government.
Exploring e-government impact on Shanghai firms' information process (Cui et al., 2006)	Survey data 1540 firms from 14 industries	Development of an integrated model to examine government factors which influence IT adoption in Chinese firms	First, the general route of IT adoption is from IT infrastructure construction to value realization. Second, government actions influence firms' IT infrastructure development and IT management decision. No evidence showing the government impact on firms' IT usage.

Title/ authors	*Research method*	*Scope of the research*	*Key conclusions/ findings*
Preface to the focus theme section: 'Electronic markets and e-government' (Pedersen et al., 2006)		Review (trajectory of electronic markets and digital government)	Electronic markets represent a viable and fruitful conceptualization of the change from bureaucratic government to digital government.
Electronic markets connecting citizens to pension reform (Ranerup, 2006)	Case study of pension reform in Sweden	Electronic markets used by public agencies as a feature in attempts to influence citizens to behave in line with public objectives (a perspective that go beyond the dominant paradigm of transactions)	Electronic markets used in pension reform in Sweden had several important objectives: to provide services in line with the 24/7 ideal for self-service; to accomplish an active choice of premium pension funds; to improve knowledge about all pension sources; and to improve the quality of the individual's choice of funds.
The effect of enacted capabilities on adoption of a government electronic procurement system by Malaysian SMEs (Salleh et al., 2006)	Survey of 206 SMEs (and 21 follow-up interviews)	The effect of enacted capabilities on adoption of a government electronic procurement system by Malaysian SMEs	Government influence was found to play an important role in the promotion or enforcement of innovative IS adoption by SMEs, and it is vital for ensuring adoption of nationwide innovative IS, particularly in developing economies. A model of how enacted capabilities affect IS adoption behavior through perceived net benefits and attitude is developed.
Electronic markets for the allocation, financing and distribution of public goods (Vragov & Kumar, 2006)	Conceptual	Concerns that obstruct the broader application of EM mechanism to the legislative branch of the government	Electronic markets have the ability to allocate, finance and distribute public goods in a decentralized manner. The use of EMs in this manner has not materialized yet due to several concerns. First, EMs may not work well because of the free-rider problem. Second, EMs can impact the status quo in a radical way and exacerbate the inequality and unfairness by differentially affecting the wealthy and the poor. Third, EMs cannot be applied to all kinds of decisions related to public goods and that they might make individual decision-making to complex.
The government of back-office integration (Bekkers, 2007)	Comparative case study	Given the political nature of back-office integration, should cross-organisational back-office be seen as a command and control challenge or a process of management challenge?	Integration is the outcome of a process in which offices have been able to create a shared understanding about the necessity of integration and in which conflicting rationalities, with their own core values, internal logic and legitimacy, have to be weighed against each other.
The myths of e-government: looking beyond the assumptions of a new and better government (Bekkers & Homburg, 2007)	Empirical research; e-government programs of AUS, CAN, UK, DK, NL	Identify, analyze and reflect on the myths underlying e-government programs	In all national policies myths of technological inevitability: a new and better government, rational information planning, and empowerment of the intelligent citizen can be discerned. The paper concludes that empirical studies have generated little support for the inescapable telos of these myths, which makes canvas cleaning effects of e-government initiatives less likely.

Title/ authors	*Research method*	*Scope of the research*	*Key conclusions/ findings*
Semantic integration of government data for water quality management (Chen, Gangopadhyay, Holden, Karabatis, & McGuire, 2007)	Study of a prototype system supporting water quality management	How can integration of data flows and business processes enable higher levels of e-government?	The paper outlines techniques to integrate numerous water quality monitoring data sources, to resolve data disparities, and to retrieve data using semantic relationships among data sources taking advantage of customized user profiles. Preliminary user feedback indicates that these techniques enhance quantity and quality of information available for water quality management.
Success factors of Geneva's e-voting system (Chevallier et al., 2007)		In-depth study of study of Geneva's e-voting system	The three main success factors of Geneva's project are: politicians were implicated at the highest level, the project built on an existing basis in the field of political rights, and the approach went step by step. Multidisciplinary teams composed of specialists in law, political rights, PR, government, security, and computer science are necessary.
Information asymmetry and information sharing (Clarkson, Jacobsen, & Batcheller, 2007)	Conceptual	Information asymmetry	A typology to classify information asymmetry into two categories, horizontal and vertical, and then identify those information sharing practices that cause the imbalance.
Varieties of interoperability in the transformation of the health-care information infrastructure (Eckman, Bennet, Kaufman, & Tenner, 2007)	IBM experience with real-world health-care information exchange engagements	Mobilization of health-care information electronically across organisations within a region or community	Current health-care information exchange is inefficient and error-prone; it is largely paper-based in most countries, fragmented, and therefore overly complex, often relying on antiquated information technology. They define the stakeholders, roles, and activities that comprise a health-care information exchange solution; they describe a spectrum of interoperability approaches and point out their advantages and disadvantages; and they look in some detail at a set of real-world scenarios, discussing the interoperability approaches that best address the needs.
Organisational solutions for overcoming barriers to e-government (Enyon & Margetts, 2007)	Online survey	Barriers to e-Government projects	Presents seven categories of barriers to e-Government progression and identifies eight associated legal areas that underpin these barriers. The discussion then turns to four organisational solutions to overcome the top barriers to e-Government as identified by an online survey of e-Government stakeholders.
A collaborative network for first responders: Lessons from the CapWIN case (Fedorowicz, Gogan, & Williams, 2007)	Case study	Examined how aspects of the external environment and the agency context impeded or facilitated the CapWIN collaborative network	A framework for aligned development of collaborative networks. Factor affecting information sharing and collaborative processes were identified as critical events, economics, politics, strategy, governance, resources, processes, and IT infrastructure.
Regional shared e-government in the region of Burgundy: the case of e-procurement (Fléri, Moutet, & Cimander, 2007)	Case study	Evaluate the service in respect to legal, functional, and technical requirements	The key factors of the success have been the governance model and the change management approach, including communication and education.

Title/ authors	*Research method*	*Scope of the research*	*Key conclusions/ findings*
RFID and interorganisational collaboration: political and administrative challenges (Gogan, Williams, & Fedorowicz, 2007)			
Maturity model mapping crime in law enforcement (Gottschalk & Tolloczko, 2007)	Conceptual	Maturity model for mapping crime in police investigations	
A general model of performance and quality for one-stop e-government service offerings (Gouscos, Kalikakis, Legal, & Papadopoulou, 2007)	Case study	Framework for solving one-stop e-government problems	One-stop government initiatives based on adoption of common standards and re-engineering of internal processes may face significant implementation risks due to a number of technical, organisational, regulatory, or political obstacles to standardization, especially in multi-national settings. An alternative approach is based on externally operating workflows that invoke internal administrative processes as they currently are.
Design governance for shared services organisations in the public services (Grant, McKnight, Uruthirapathy, & Brown, 2007)		Governance design for shared services organisations in the public sector	Strategic and operational concerns should be separate. The design of shred services is most effective when those who deliver services to clients are separate from those who ensure compliance with corporate policy and standards.
Business interoperability profiles (Greiner, Legner, Lippe, & Wende, 2007)	Conceptual	Modeling of cross-organisational business processes (CBP)	The successful implementation of CBPs requires a clear understanding of the common processes across all involved stakeholders. They developed a framework to support modeling of CBPs.
Business interoperability research: present achievements and upcoming challenges (Legner & Lebreton, 2007)		Review	Interoperability research is at an early stage and the definitions and concepts underlying business interoperability are still under discussion. Theoretical concepts and models which are empirically validated are still highly underrepresented in the interoperability research.
Collaborative digital government in Mexico: Some lessons from federal web-based interorganisational information integration initiatives (Luna-Reyes et al., 2007)	Case study in the Mexican federal government	Digital government in Mexico in terms of federal web-based interorganisational information integration initiatives	Institutional arrangements and organisational structures shape not only the enacted technology, but also other process and results of government IT projects.
Interoperability of e-government information systems: issues of identification and data sharing (Otjacques, Hitzelberger, & Feltz, 2007)	Exploratory study, status in 18 EU member states	The way public organisations manage identity-related data and the sharing of such data	Despite significant progress in harmonizing the legal and administrative provisions and technical standards in the European Union, there are still considerable cross-country differences regarding collection and sharing of identity-data.

Title/ authors	*Research method*	*Scope of the research*	*Key conclusions/ findings*
Interorganisational information integration: a key enabler for digital government (Pardo & Tayi, 2007)			
Electronic Government as a combination of human and technological agency: Testing the principle of symmetry (Ranerup, 2007)	Case study	Computerized decision support for portfolio management in the premium pension system in Sweden	A system incorporating humans in conjunction with DSS is able to make advanced economic decisions in line with pension policy by incorporating the preferences and characteristics of humans as well as the capabilities of technology.
Setting interoperability standards for e-government: an exploratory case study (Santos & Reinhard, 2007)			
E-Government integration and interoperability: framing the research agenda (Scholl & Klischewski, 2007)		Overview of the e-government literature on integration, interoperation, interoperability, and information sharing	Future research framework revolves around the proposed research directions of: 1) foci and purposes, 2) limitations and constraints, and 3) process and outcomes.
Towards end-to-end government performance management: case study of interorganisational information integration in emergency medical services (EMS) (Schooley & Horan, 2007)	Case study	Interorganisational information sharing in the public domain	The case study illustrates promising factors that can enhance information sharing across organisations, while noting that considerable gaps remain in achieving an end-to-end IT-enabled performance approach.
Urban information integration for advanced e-planning in Europe (Wang, Song, Hamilton, & Curwell, 2007)	Case study	Technical approaches for multidimensional information integration	The achievements and further development of multidimensional information integration through the use of innovative urban data modelling techniques are discussed
Interoperability and the exchange of good practice cases (Archmann & Kudlacek, 2008)		Governance aspect of interoperability	The governance aspect of interoperability includes four type of key factors: political, legal, managerial, and economic.
A comprehensive framework for the assessment of eGovernment projects (Esteves & Joseph, 2008)	Conceptual	Framework for the assessment of e-Government initiatives	The three dimensions are e-Government maturity level, stakeholders, and assessment levels.
Stages of e-government interoperability (Gottschalk & Solli-Sæther, 2008)	Conceptual	Maturity model of e-government interoperability	Improved interoperability between public organisations as well as between public and private organisations is of critical importance to make electronic government more successful. In this paper, stages of e-government interoperability is identified and discussed. Four stages are presented: work process stage, knowledge sharing stage, value creation stage, and strategy alignment stage.

Title/ authors	*Research method*	*Scope of the research*	*Key conclusions/ findings*
Information quality as capstone in negotiating e-government integration, interoperation and information sharing (Klischewski & Scholl, 2008)			
Collaborative management and e-government: a survey of state government CIOs (Reddick, 2008)			

The review included the following major journals in the e-Government area, up to the most current volume available by August 2008; *Electronic Government, Electronic Markets, European Journal of ePractice, Government Information Quarterly, Public Management Review.* Promising articles were followed back to their origin, whether based in articles, books, or dissertations.

Part III

Semantic interoperability

17 THE CHALLENGE OF SEMANTIC INTEROPERABILITY

The quest for semantic interoperability arises in situations in which there is need for sharing or exchange of data (or information), and where it is not crystal clear to the intended users how the shared or exchanged data are to be interpreted. For such an enterprise to be successful one must at least be able to transmit (or share)

- a description of the *meaning* of the data which is accurately enough to enable correct use of the data on the receiver's side
- the data (or *records*) themselves

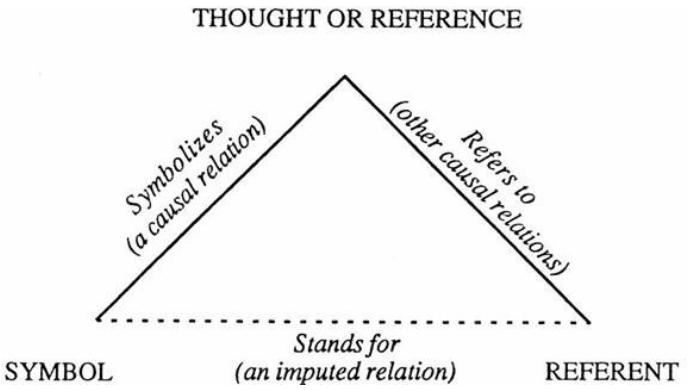
We shall first give an introduction to the way in which questions related to meaning are approached in this report. Then we shall address questions related to data such as information governance. Finally we shall address the notion of semantic interoperability and put it into a wider context, using an interoperability framework from simulation theory.

17.1 Some clarifications about semantics

When addressing questions of interoperability and meaning, we should first bear in mind that information that we want to exchange is always *about* something, about things that are, in the widest sense of the word, objects in the real world. They can, to name a few, be concrete objects (like persons), abstract objects (like a particular salary), concepts (like "salary") and tokens (like 'salary'). Second, the information ascribes properties to objects (like salary to a person). Semantic interoperability is about, first, being able to *refer unambiguously* to the objects about which we have information and, second, about being able to identify *what properties* we ascribe to them when we represent information.

The first task is solved if we have a unique *name* for each object. In some rare cases names follow conventions or commonly accepted standards, which is the case when person number is used to identify persons. In most situations, however, tokens function as names only in specific contexts. Once we step out of the context we thus have to further qualify the names to preserve uniqueness of reference, i.e., we have to equip the tokens with so-called *identity criteria*. We shall address the question of names and identity in the next section.

The exchange of properties ascribed to objects is more subtle and calls for some initial clarification. There is a very long tradition in the history of philosophy to use a three-way distinction when discussing matters related to the meaning of signs and words. In an influential work from 1923 Ogden and Richards [Ogden] drew a simple picture of a three-way distinction of this kind, which is now known as the *semiotic triangle*. This is how it looks in its original form:



Broadly speaking, the relationship between symbols, concepts and referents, according to the semiotic triangle, may be summarized as follows:

- The *symbol* is a token of an abstract type which is used to *denote* objects (or set of objects) in the real world. A symbol could be a word in a language, characters, digits, a traffic sign or an icon on your computer desktop. The dashed line in the triangle indicates that the relation of reference is indirect; a symbol succeeds in referring through some form of mental activity of, say, the speaker and the addressee.
- The *reference* is an abstract, atemporal unit that may be taken as the information content of symbols and sentences. It is something that can be communicated between speakers that are possibly using different words (or languages) to convey it.
- The *referent* is something that a concept refers to; it can, e.g., be abstract or concrete, material or immaterial, as pointed out above.

The referent of the concept "tree" is (in its most basic sense) simply the set of all trees (where "all" is taken in its widest sense). While the original triangle aims at characterizing human communication in general, it has in the context of semantic interoperability a more restricted application. Clearly, semantic interoperability is a special instance of human communication, and it is for the purpose of this report fruitful to qualify the use of triangle accordingly.

To prevent common misconceptions, let us initially point out a few things that we shall *not* be addressing. We shall not be concerned about what kind of existence it is that referents have. This is a metaphysical issue; all that we need to assume about the objects is that we have names for them. And we shall not enter into discussion about what kind of entities that information contents are, and how they relate to human mind; these questions belong to the philosophy of mind and metaphysics, and have no direct relevance for our discussion here. In fact we shall ignore all mental aspects that pertain to information content.

Instead we shall in a restricted sense be talking about *concepts.* We will assume that a concept has a *definition,* and will in most cases not have to distinguish between a concept and its definition. We will say that a concept can be predicated of a thing, and that it is either true or false of any given object; if true we say that the concept *subsumes* the object. This leads to a somewhat simpler view:

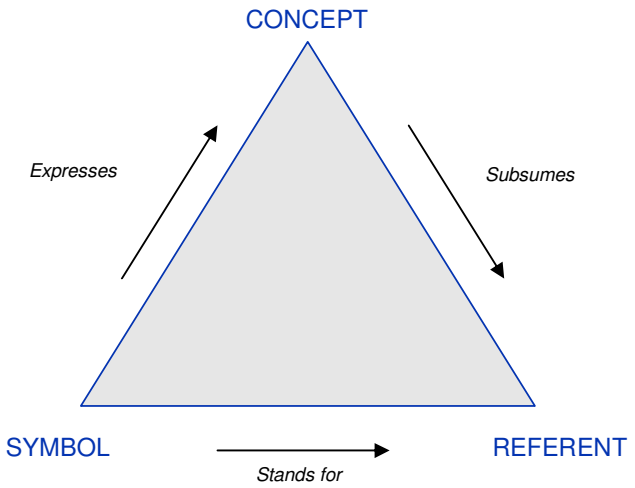


Figure 1: The semiotic triangle

The triangle can be used to make three points that are important for semantic interoperability.

- The same symbol can denote different things. When we use the word 'tree' we will not achieve interoperability unless it is clear to all parties what sense of the word we are using.
- Different symbols can be used to denote the same thing. A system using Norwegian terms will not interoperate with a system using English terms if we are not able to link the symbol 'tre' as used in the former with the symbol 'tree' as used in the latter.
- To achieve interoperability we need to fix a language in which we can express concepts (with unique definitions) unambiguously. If we have such a language at our disposition it suffices for interoperability to exchange the concepts along with the data.

The figure below shows different meanings (pointers to different concepts) of the word 'tree'. One of the definitions of the symbol "tree" is defined in Webster online as: "a tall perennial woody plant having a main trunk…" [16] A definition of this kind identifies a concept. If we agree on the definition of this concept and of the name, say T, of a specific wooden thing in my garden, I can say of T that it is a tree and be sure that the correct message gets across when we exchange information.

[16] Merriam Webster http://www.m-w.com/dictionary/tree

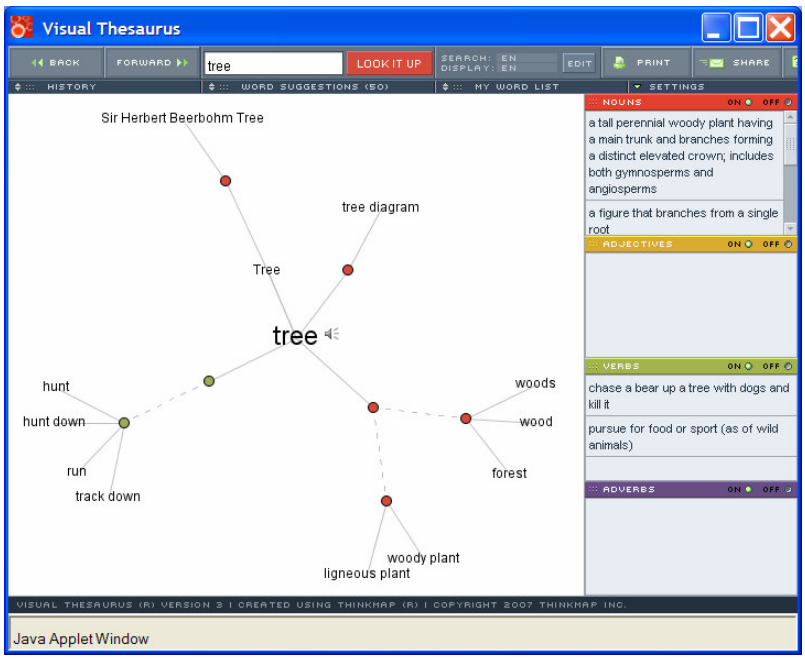


Figure 1 visualizing relations of words

An important function of semantic technologies is to support the construction of systems of concepts and their interrelations. Such systems are called *terminologies* or *ontologies*. The term 'ontology' means in this context nothing else than vocabularies consisting of terms (i.e. concepts) and relationship between them, typically defined through definitions, implications and binary relations. It should not be confused with philosophical ontology, which is just another name for metaphysics.

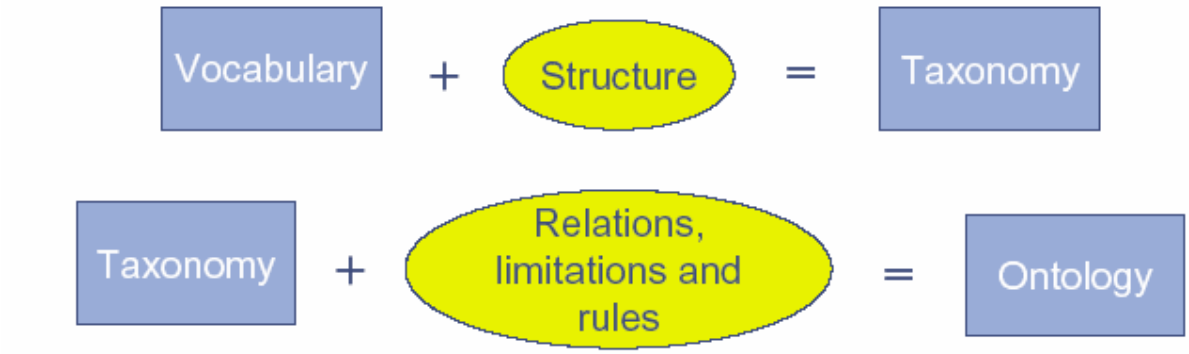


Figure from [FFI]. The degree of formalization gives different terminological structures.

The definitions that constitute a terminological system can be more or less formalized. The lowest level of formalization is just a dictionary with definitions in natural language. If we formalize subclass relations and add this structure to a dictionary we have the basic building block of a taxonomy. In an ontology the definitions have a very high level of formalization, i.e. a level which makes them machine manipulable. This opens for the advanced functionality of reasoning services, with potentially very powerful applications. However, when the degree of formalization increases, so does the

complexity of constructing the ontologies, and also the cost. Ontologies and ontology languages are further addressed in the next section.

For a terminology to be called an ontology it is usual to require that its definitions are accepted among a reasonably large group, i.e. it shall not be private. It should also include a specification of the meaning of the terms in the vocabulary which constrains the interpretation of the terms.
The term 'ontology' is most often associated with W3C recommendations (OWL in particular) and work in this tradition. The W3C approach is in many ways complementary to OMG-based approaches to interoperability; model-based solutions and ontology-based solutions can be viewed as two approaches to achieving semantic interoperability. While one in a model-driven approach aims at defining a semantic mapping between two models directly, one will in an ontology approach design an external vocabulary, which one will connect to the individual models by means of wrappers. Roughly, the ontology approach offer better support for maintaining and constructing large ontologies with non-trivial structure, while the OMG approach has better support for graphical visualization and integration within a general architecture. However, OMG-based tools offer little reasoning support beyond simple inheritance.

There is currently active research that attempts to link the W3C and OMG approaches to interoperability. It is likely that this question will be addressed in depth by Semicolon; discussion of this important question, and in particular the extent to which it has a bearing on already ongoing activities of Semicolon partners, requires research and is therefore scheduled for further Semicolon work in the near future.

17.2 Semantic and information interoperability

As defined in Wikipedia, Semantic Interoperability[17] (also referred to as Computable Semantic Interoperability) is *the ability of two or more computer systems to exchange information and have the meaning of that information automatically interpreted by the receiving system accurately enough to produce useful results, as defined by the end users of both systems.* The current review of Semantic Interoperability is limited to issues of interpretation and meaning in information exchange between information systems, such as models and vocabularies for decision support and command & control systems. As noted by several authors, formulation and interpretation of meaning in organizational, social and political contexts are often even more important than the merely technical and informational issues.

According to the IEEE definition of interoperability of information systems we define *Information Interoperability (II)* [Information Interoperability] as *the ability of two or more systems or components to exchange information and to use the information that has been exchanged*. With this definition we can define a model of Information Interoperability:

[17] http://en.wikipedia.org/wiki/Semantic_interoperability

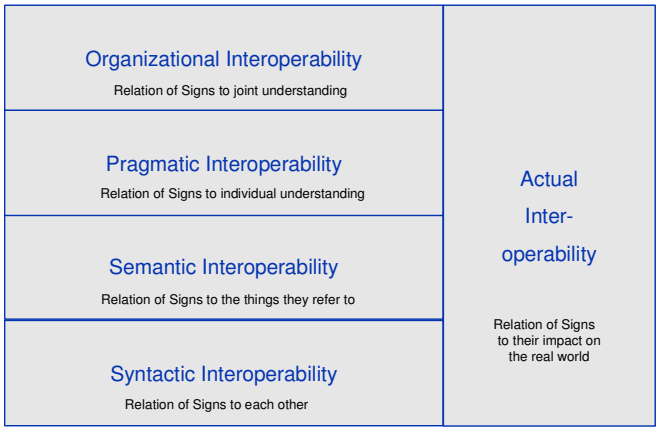


Figure 1: An Information Interoperability (II) Model

In this context we regard signs (or symbols) as the collection of tokens (letters) we use to create messages (write or say) when we communicate (exchange information), and –

- *Syntactic Interoperability* to mean that what you hear is the same as what I say.
- *Semantic Interoperability* to mean that what you understand from what you hear is the same as what I mean by what I say.
- *Pragmatic Interoperability* to mean that what you can do with what you understand is the same as what I want to achieve with what I mean.
- *Organizational Interoperability* to mean that what we can do together with what we say and hear is the same as what we mean and understand to be our common goals.
- *Actual Interoperability* to mean that what we actually do with what we say and hear is related to what we actually want to achieve.

Clearly these levels of Interoperability are related, and dependent on each other. For example, collective Organizational Interoperability requires shared understanding through individual Pragmatic Interoperability (if we shall understand each other as a group, we also have to understand each other as individuals). Likewise, it is not possible to achieve Actual Interoperability in an organization without Syntactic, Semantic, Pragmatic and Organizational interoperability (if we do not understand each other at all levels we can not cooperate and collaborate as required to reach common goals).

Following recent work in simulation theory we discuss an updated version of the Levels of Conceptual Interoperability [Conceptual Interoperability] Model (LCIM) [Tolk] as a basis for defining a Model of Information Interoperability. Extended versions of LCIM [Turnitsa] take a broad view of information exchange and define the following levels of interoperability –

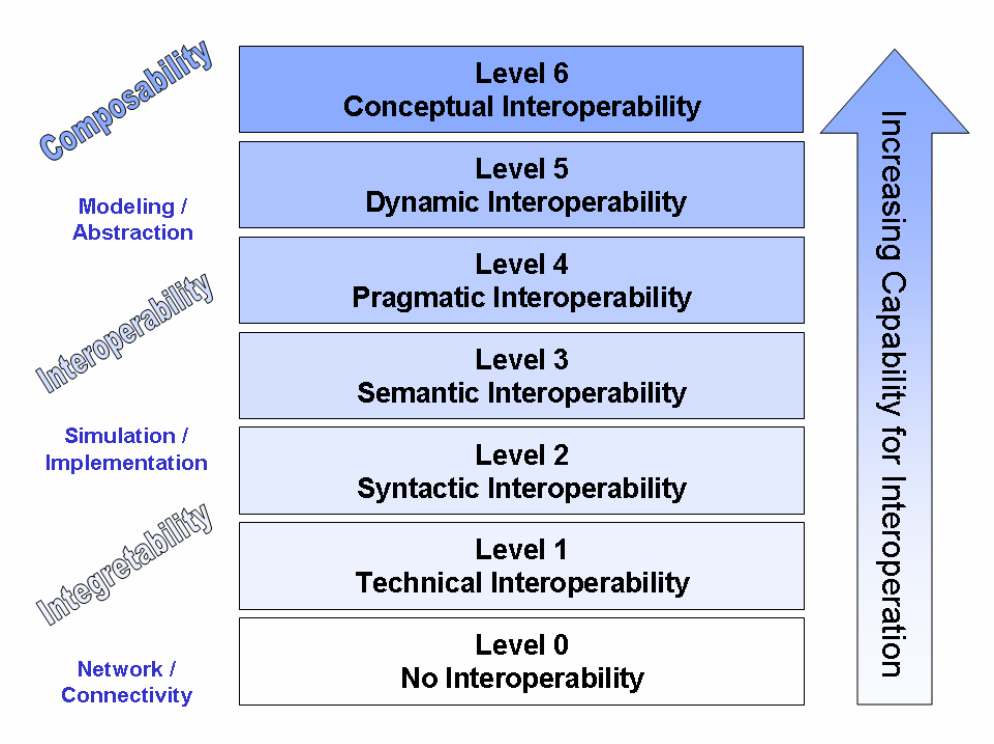


Figure 2: The Levels of Conceptual Interoperability Model (LCIM)

In figure 1 the levels of Interoperability are defined [Turnitsa] as follows -

- Level 0: Stand-alone systems have No Interoperability.
- Level 1: On the level of Technical Interoperability, a communication protocol exists for exchanging data between participating systems. On this level, a communication infrastructure is established allowing systems to exchange bits and bytes, and the underlying networks and protocols are unambiguously defined.
- Level 2: The Syntactic Interoperability level introduces a common structure to exchange information; i.e., a common data format is applied. On this level, a common protocol to structure the data is used; the format of the information exchange is unambiguously defined. This layer defines structure.
- Level 3: If a common information exchange reference model is used, the level of Semantic Interoperability is reached. On this level, the meaning of the data is shared; the content of the information exchange requests are unambiguously defined. This layer defines (word) meaning. There is a related but slightly different interpretation of the phrase Semantic Interoperability which is closer to what is here termed *Conceptual Interoperability*, i.e. information in a form whose meaning is independent of the application generating or using it.
- Level 4: Pragmatic Interoperability is reached when the interoperating systems are aware of the methods and procedures that each system is employing. In other words, the use of the data – or the context of its application – is understood by the participating systems; the context in which the information is exchanged is unambiguously defined. This layer puts the (word) meaning into context.

- Level 5: As a system operates on data over time, the state of that system will change, and this includes the assumptions and constraints that affect its data interchange. If systems have attained Dynamic Interoperability, they are able to comprehend the state changes that occur in the assumptions and constraints that each is making over time, and they are able to take advantage of those changes. When interested specifically in the effects of operations, this becomes increasingly important; the effect of the information exchange within the participating systems is unambiguously defined.
- Level 6: Finally, if the conceptual model – i.e. the assumptions and constraints of the meaningful abstraction of reality – are aligned, the highest level of interoperability is reached: Conceptual Interoperability. This requires that conceptual models are documented based on engineering methods enabling their interpretation and evaluation by other engineers. In essence, this requires a "fully specified, but implementation independent model" as requested by Davis and Anderson [Turnitsa]; this is not simply text describing the conceptual idea.

The LCIM focuses on technical support by information systems, such as command and control information systems in the military context. As Alberts and Hayes point out in [Alberts], the organizational and social aspects are often even more important. In [Conceptual Interoperability] Tolk proposes such a layered framework for measures of merits dealing with questions like tactical or strategic alignment of objectives or even political will of coalition partners in.

To complement the LCIM, Page et al. [Page] suggest defining *composability* as the realm of the model and *interoperability* as the realm of the software implementation of the model. In addition, their research introduces *integratability* coping with the hardware-side and configuration side of connectivity -

- Integratability contends with the physical/ technical realms of connections between systems, which include hardware and firmware, protocols, etc.
- Interoperability contends with the software- and implementation details of interoperations, including exchange of data elements based on a common data interpretation, etc.
- Composability contends with the alignment of issues on the modeling level. The underlying models are purposeful abstractions of reality used for the conceptualization being implemented by the resulting simulation systems.

The LCIM has been successfully applied not only in the domain of Modeling & Simulation, but generally to assess a range of model-based interoperability challenges [Davis], [Alberts] and [Tolk2003].

18 FROM TERMS TO MODELS OF SEMANTICS

In this chapter we first present the approach to ontologies taken by W3C. We then address some issues with a more general scope like methodologies for building ontologies, how they change and evolve and some questions related to ontology quality. Of particular relevance for Semicolon is how ontologies can be used in data integration; this is discussed in a separate section as an illustration of an application. Finally we address the question of how identities of concepts, referents and information of referents may be handled.

Issues like how to establish routines, organisational bodies and technical solutions for maintaining common ontologies centrally or distributed organised are also mentioned in the following chapters.

18.1 Ontologies and ontology languages: the W3C perspective

Ontologies are formalised vocabularies of terms covering specific subjects like human anatomy and technical equipment in process plant industries. The typically contain definitions of terms through relationships with other terms of the ontology. In an ontology for anatomy we may for instance define a Heart as a Muscular Organ that is a part of a Circulatory System. Ontologies are often divided into an upper part comprised of general terms like the "part of" relation and the most basic categories, and domain-specific parts covering the concepts that are particular to the area of the ontology.

Ontologies have proved very useful in a number of domains; in life sciences there are for instance ontologies in use with more than a hundred thousand concepts. As maintaining consistency for even quite small ontologies is difficult, automated reasoning is used to detect inconsistencies in the concept definitions of ontologies. Availability of tool support of this type is particularly strong for the W3C recommendation OWL, a fact which has made OWL the language of choice for the design of most ontologies today. In OWL one can express simple relationships like "A man is a person" and "Bill is married to Ann," but also more complex definitions like "A bachelor is a man that is not married"; from these statements one can e.g. draw the conclusion that "Bill is not a bachelor" by means of valid inference patterns of *Description logic* that underlies the definition of OWL.

W3C has above OWL introduced the rule language SWRL that can express more than OWL, and which will be considered in this project. A fundamental challenge with both OWL and SWRL is the inherent computational complexity of the languages. However, state of the art reasoning engines are much optimised and are able to cope efficiently with large ontologies where more naïve implementations would be hopelessly intractable.

The discipline of knowledge representation, generally, aims at developing abstract formalisms that are sufficiently expressive to be used as a basis for intelligent applications. In this context 'intelligence' is understood as the ability of a system to reason about a domain by tracing the consequences of an explicit representation of that domain, thereby facilitating efficient information management such as retrieval, updating and exchange.

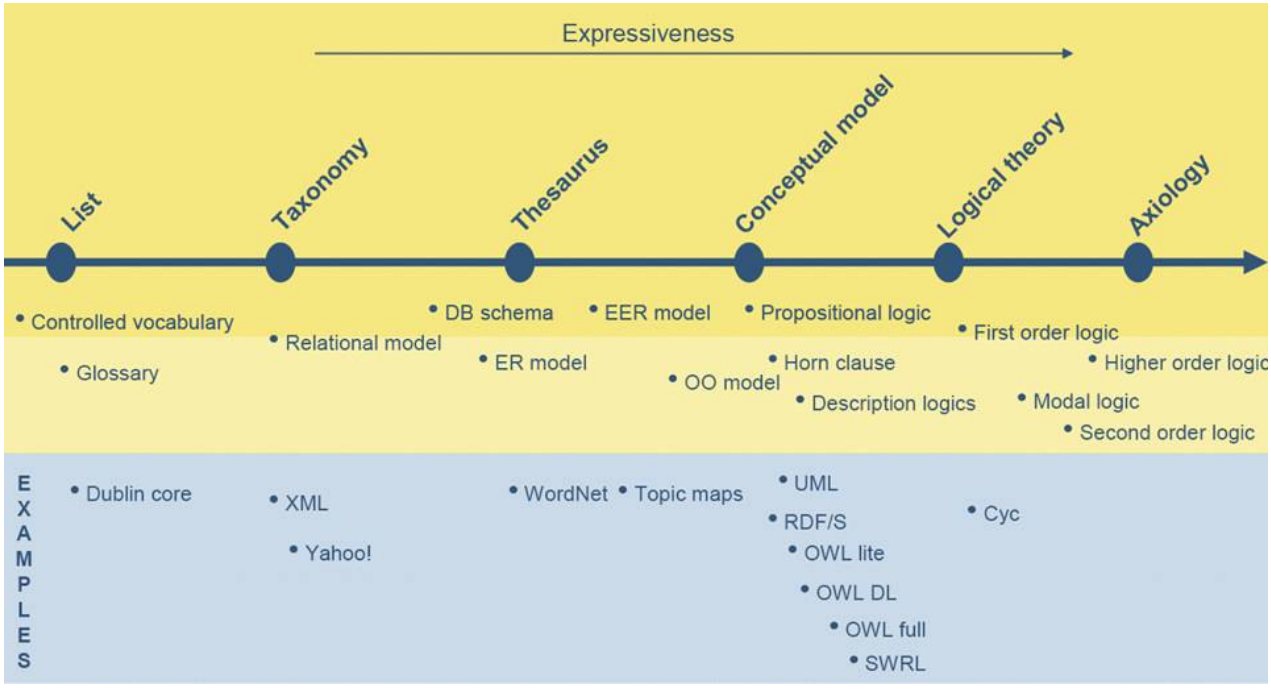
Description logics are a family of languages for classifying concepts and individuals *per genus et specificam*. Stated differently, a description logic is a language for building *taxonomies* or subsumption hierarchies that provide information about the logical interrelations between a set of concepts and their instances. Classification of individuals thus determines whether an individual exemplifies a certain concept, that is, whether an instance relationship is implied by the description of the individual and the definition of the concept. Depending on the depth of a concept in the subsumption hierarchy, a taxonomy thus offers views on the data on different levels of abstractness. It can therefore be used to speed up inference services that may not require a low level of granularity, as well as services that exploit the commonalities shared by a set of objects.

Unlike many other knowledge representation formalism, such as for instance semantic networks, frame systems, conceptual graphs and entity-relationship diagrams, description logics are *logics*. That is, they are calculi equipped with a formal semantics that allows the consequences of a set of assertions to be determined with precision. This, in turn, is a prerequisite for determining the computational properties (such as decidability and complexity) of the reasoning services they are meant to provide. Research on description logics covers mathematical foundations as well as actual implementations of knowledge representation systems. Several different application areas have been explored, for instance

- Natural language processing, such as the VERBMOBIL and LILOG projects.
- Configuration and manufacturing technology, e. g. the PROSE and QUESTAR product line.
- Software engineering, for instance the CODEBASE system.
- Medical informatics, e.g. the GALEN server, and the medSYNDICATE natural language processor for medical texts.
- Digital libraries, for instance the FindUR system.

Many of these applications have been relatively successful. This is in large part due to the methodology adopted, which has been based on a very close interaction between theory and practice [Nardi]. On the one hand there are various implemented systems based on description logics that highlight the requirements that a workable and useful knowledge-management system should meet. On the other hand there is the mathematical research into the computational properties of reasoning. These properties have been mapped out in detail for most of the different logics on offer. The trade-off between expressiveness and tractability is particularly well understood (a survey of basic results can be found in [Doninini]), and allows software designers to select language constructs in a modular fashion whilst staying below the threshold of combinatorial explosion. Several general-purpose reasoning engines exist that generate classification hierarchies from a description of data, and implement functionality such as calculation of least common subsumers and most specific superconcepts. Notable examples are FACT++, KAON2 and RACER. It is a mark of progress that these different technologies are now beginning to fuse and integrate. In particular, a lot of activity is currently clustering around W3C's web ontology language OWL, which is itself an XML port of (variants of) description logic. A notable example is the Protégé editor. Protégé is a free open-source ontology editor and a knowledge-base framework, which comes pre-packaged with the FACT++ reasoning engine, provides a plugin for RACER and exports ontologies to OWL.

A general point to note is that there is a systematic connection between the expressivity of the ontology language and the complexity of the key inference problems that are associated with the language. The description logic that underlies OWL is decidable, which means that every question of logical consequence can be determined in finite time, while still very expressive. The figure below puts the expressivity in relation to other well-known representation formalisms.



Source: Davis, M. 2006, Semantic Wave 2006 [Davis M]

Application of ontologies include intelligent search and query handling and data integration, both of which will be explored in Semicolon. Ontology-based data integration is addressed further below.

18.2 Topics from ontology engineering

This section is based on the chapters "Ontology Evolution", "Ontology Mediation, Merging and Alignment" and "Ontology Engineering methodologies" in [SemTechTrends] and the chapters "Theoretical Foundation of Ontologies", "The most outstanding Ontologies" and "Methodologies and Methods for Building Ontologies" in [Perez]. The chapter is more an overview of different directions of research activities and known challenges than what the actual state of the art of what problems are solved and how.

In public sector there will exist many ontologies maybe one for each public agency or even each division within each agency. Based on this background there is a large need to know what we can do with operations like, union, intersection, different types of comparing, extracting etc. And these operations need to be done even if the different ontologies are made by different methodologies and are represented in different representation formats.

Separating the knowledge model / ontology from software and data, brings new opportunities into dynamic behaviour and process change. Technology support in these processes and the ability to do effective software development and maintenance can be improved. This ability is of great interest for both the record repositories and the software used when handling preservation of semantic value.

18.2.1 Ontology engineering methodologies

Ontologies aim to capture consensual knowledge in a generic way, and are meant to be reused and shared across software applications and by groups of people. They are usually built cooperatively by different groups of people in different locations [Perez]. From library science we have a tradition of systematically classifying records by thesauri, and a set of approaches has been developed [Gunnlaugsdottir].

In the software system development tradition, most models of data in databases have for the last 20 years been made using entity relationship models. Object-oriented programming languages have, as part of the software models, made models of the data used in the software. Ontology engineering models the knowledge or the meaning of data, and not the way data is stored, exchanged or represented.

Ontology engineering, the way we make ontologies, is a hot research topic with many variations and many different scopes and targets. We separate the domain expert having the knowledge we would like to model, from the ontology engineers who know how to capture the knowledge in ontology models. Separate tools are often used for capturing and maintaining ontologies.

An important goal of ontologies is to separate the formal model of knowledge from the records and the software used. The software tools available to average software developers still lack much of the ability to gain from this separation. This is a challenge leading to slow uptake of semantic technology.

Ontologies may give improved usability aspects related to searching, quality checking of model and data, navigation, data mining, architecture layering, reuse of data, exchange of data etc. Further we will look into different aspects of development, maintenance and support activities related to ontology engineering.

- Ontology management activities are to define/describe: Control mechanism, schedules, responsibilities, quality steps etc, maintenance and evolution procedures. These activities are similar to general activities in projects for developing systems or software of some kind.
- Ontology development activities are: Do feasibility studies, define environment/framework where the resulting ontology should be used, choose and describe methodology for how to conceptualize, formalize, represent, and implement the target ontology, describe regime for ontology population, use etc.
- Ontology support activities are: Knowledge acquisition, evaluation, system integration, ontology merging and alignment, and configuration management.

The ontology engineering methodology has similarities with the software development approach called the water fall model. Ontology evolution has similarities to iterative software development.

Further relevant research is to look into topics like:

- Framework for comparing ontology engineering methodologies, e.g. [Cuel] “A survey on ontology creation methodologies”.
- Diligent methodology, distributed ontology engineering and frequently changing user needs.
- Ontology evaluation, e.g. OntoClean [Guarino]
- Quality metrics for ontologies described in a separate chapter below.

18.2.2 Ontology quality

The article [Barry Smith] "A Realism-Based Approach to the Evolution of Biomedical Ontologies" discusses ways to calculate the quality of ontology evolution. There are initiatives like OntoClean to develop methodologies for validating the ontological adequacy of taxonomic relationships. OntoClean has provided a logical basis for arguing against the most common modelling pitfalls, and argues for what we have called "clean ontologies" [Guarino].

18.2.3 Knowledge base evolution

In the field of knowledge management [Gunnlaugsdottir], the transition from personal knowledge to public and documented knowledge is a circular, continuous movement. Personal knowledge is based on public and/or some other persons' knowledge. When new personal knowledge is documented it can become published, public and then be a part of a common knowledge base. In the Knowledge Management Field, knowledge can be recorded in records and made public available.

18.2.4 Ontology distribution and distributed knowledge base

The knowledge base consists of the information stored in the records themselves and in the relevant ontologies. Additionally, referral may be made to records maintained by external parties. Thus, one may need to rely on the continued support and trustworthiness of external parties and their record management regimes. This situation is relevant to both distributed ontologies and record archives which are not self contained. Linkage between records located in different locations will be challenging to maintain.

18.2.5 Ontology evolution

During its lifetime an ontology will be changed by its users and the changing content in the base repository. This topic is of high relevance to Semicolon partners in light of ontology- governances and life time issues.

The two main types of changes in an ontology are *usage-driven* and *data-driven* changes.

- Usage-driven: there is a change or unbalance in how users use the archive/knowledge portal, and what support the portal can have from its ontology(ies). The use of records have changed, and the ontology needs to be changed accordingly. As input to the ontology change process, use of pattern recognition for tracking user behaviour or explicit changes in work procedures could be useful. Also search logs and analysis could be a source for capturing usage-driven changes.
- Data-driven means that the records contained in the digital repository have changed properties, volume or the new records are not good enough reflected in the ontology. Techniques from e.g. data mining and business intelligence are used to discover data-driven changes. And the results of using these techniques will be useful input to the change process.

Knowledge and usage of records change. If those maintaining an ontology do not manage to reflect the changes and fail to have the ontology updated, the ontology will not serve its purpose in a semantic solution. Ontology evolution starts with identifying changes.

Based on [SemTechTrends, p. 52] one way of illustrating the stages in ontology evolution is as follows:

1. Identifying changes
2. Representation of changes
3. Semantics of changes
4. Implementation of changes
5. Propagation (check if dependent ontologies, artefacts and systems are consistent after changes.)
6. Validation of changes

From Validation the circle goes back to Capturing.

18.2.6 Ontology mediation

Ontology mediation is used to share data between heterogeneous sources, and/or to reuse data from different knowledge bases. Established approaches for ontology mediation are ontology-mismatches, mapping, alignment and merging. As explained in the chapters below, the methods have some overlap both in what kind of problem they are meant for and the properties of outcome.

Different types of ontology mismatches are shortly described as follows:

- Conceptualization mismatch
  - *Scope mismatch* occurs when e.g. one ontology describes patients and the other describes taxpayers. The referents they refer to overlap but are different.
  - *Granularity level mismatch*, e.g. if one ontology describes persons at most detailed level and the other ontology describes subcategories of persons.
  - *Different worldviews* of a reference. Depending on the worldview, the ontology reflecting a reference with the corresponding referents could be described differently. E.g. there may be different views of the country Palestine, its borders, its existence etc., depending on who is making and maintaining the ontology.
- Different ways of describing a reference (explication mismatch)
  - Modelling style/paradigm,
    - use of attributes versus sub-classes
    - use of point in-time versus intervals
  - Terminology mismatch
    - Identical reference in the ontologies, but the references are linked to different symbols (synonym problem).
  - Encoding mismatch occurs when one ontology uses e.g. meters for a certain reference and another uses feet or centimetres.

Ontology mapping takes two or more ontologies as input, and makes a new, separate ontology describing the mappings (bridge) between the sources. The source ontologies continue their separate life.

Ontology alignment is the process of discovering similarities between two source ontologies. Input is the relevant ontologies, and output of the process is a specification of the correspondence between the ontologies.

Ontology merging takes two or more ontologies as input, and makes a new ontology which is the union of the sources, plus the needed links between the sources. Some of the methods for ontology merging keep the original ontologies as is, and add a bridge ontology. The source ontologies and the bridge ontology are used together as a whole in further operations like record translation or querying.

18.3 Upper ontologies

In information science, an upper ontology (top-level ontology or foundation ontology) is an attempt to create an ontology which describes very general concepts that are the same across all domains. The aim is very broad semantic interoperability between large numbers of ontologies accessible "under" this upper ontology. As the metaphor suggests, it is usually a hierarchy of entities and associated rules (both theorems and regulations) that attempts to describe those general entities that do not belong to a specific problem domain.

A well-known and quite comprehensive ontology available today is Cyc, a proprietary system under development since 1985, consisting of a foundation ontology and several domain-specific ontologies (called *micro-theories*). A subset of that ontology has been released for free under the name OpenCyc, and a more or less unabridged version is made available for non-commercial use under the name ResearchCyc.

The ISO 15926 (POSC Caesar) standard [ISO 15926] is titled: "*Industrial automation systems and integration—Integration of life-cycle data for process plants including oil and gas production facilities*" is a standard for data integration, sharing, exchange, and hand-over between computer systems. This title is regarded too narrow by the present ISO 15926 developers. Having developed a generic data model and Reference Data Library for process plants it turned out that this subject is already so wide, that actually any state information may be modeled with it.

ISO 15926 has 7 parts:

- Part 1 - Introduction, information concerning engineering, construction and operation of production facilities is created, used and modified by many different organizations throughout a facility's lifetime.
- Part 2 Data Model. a generic 4D model that can support all disciplines, supply chain company types and life cycle stages, regarding information about functional requirements, physical solutions, types of objects and individual objects as well as activities.
- Part 3 - Geometry and Topology, defining, in OWL, the geometrical constructs of ISO 10303-42.
- Parts 4 [ISO 15926-RDS] - Reference Data, the terms used within facilities for the process industry.
- Parts 5 and 6 – Procedures for adding new Reference Data and maintaining the Reference Data Library.
- Part 7 [ISO 15926-Templates] - Implementation methods for the integration of distributed systems, defining an implementation architecture that is based on the W3C Recommendations for the Semantic Web.

The BFO or Basic Formal Ontology framework developed by Barry Smith and his associates consists in a series of sub-ontologies at different levels of granularity. The ontologies are divided into two

varieties: SNAP (or snapshot) ontologies, comprehending continuant entities such as three-dimensional enduring objects, and SPAN ontologies, comprehending processes conceived as extended through (or as spanning) time. BFO thus incorporates both three-dimensionalist and four-dimensionalist perspectives on reality within a single framework. Interrelations are defined between the two types of ontologies in a way which gives BFO the facility to deal with both static/spatial and dynamic/temporal features of reality. Each SNAP ontology is an inventory of all entities existing at a time. Each SPAN ontology is an inventory (processory) of all the processes unfolding through a given interval of time. Both types of ontology serve as basis for a series of sub-ontologies, each of which can be conceived as a window on a certain portion of reality at a given level of granularity. An example of an application of BFO can be seen in the Ontology for Biomedical Investigations (OBI).

Developed by Nicola Guarino and his associates at the Laboratory for Applied Ontology (LOA), the Descriptive Ontology for Linguistic and Cognitive Engineering (DOLCE) is the first module of the WonderWeb foundational ontologies library. As implied by its acronym, DOLCE has a clear *cognitive bias*, in that it aims at capturing the ontological categories underlying natural language and human commonsense. DOLCE, however, does not commit to a strictly referential metaphysics related to the intrinsic nature of the world. Rather, the categories it introduces are thought of as cognitive artifacts, which are ultimately depending on human perception, cultural imprints and social conventions. In this sense, they intend to be *descriptive* (vs *prescriptive*) notions that assist in making already formed conceptualizations explicit. DOLCE is an ontology of particulars, in the sense that its domain of discourse is restricted to them. Of course, universals are used to organize and characterize the particulars, but are not themselves subject to being organized and characterized (as meta-properties).

DnS (Descriptions and Situations), developed by Aldo Gangemi (LOA, Rome), is a *constructivist* ontology that pushes DOLCE's descriptive stance even further. DnS does not put restrictions on the type of entities and relations that one may want to postulate, either as a domain specification, or as an upper ontology, and it allows for context-sensitive '*redescriptions*' of the types and relations postulated by other given ontologies (or 'ground' vocabularies). The current OWL encoding of DnS assumes DOLCE as a ground top-level vocabulary. DnS and related modules also exploit 'Codeps' (Content Ontology Design Patterns), a newly created tool which provides a framework to annotate 'focused' fragments of a reference ontology (i.e., the parts of an ontology containing the types and relations that underlie 'expert reasoning' in given fields or communities).

Both DOLCE and DnS are particularly devoted to the treatment of social entities, such as e.g. organizations, collectives, plans, norms, and information objects. The DOLCE-2.1-Lite-Plus OWL version, including a number of DnS-based modules, has been and is being applied to several ontology projects. A lighter OWL axiomatization of DOLCE and DnS, which also simplifies the names of many classes and properties, adds extensive inline comments, and thoroughly aligns to the repository of Content patterns (available soon as a collaborative design portal) is now available as DOLCE-Ultralite (abbreviated: DUL). Despite its simplification, which greatly speeds up consistency checking and classification of OWL domain ontologies that are plugged to it, the expressivity of DOLCE-Ultralite is not significantly different from the previous DOLCE-Lite-Plus. DOLCE OWL versions, DOLCE-Ultralite and the pattern repository are developed and maintained by Aldo Gangemi [Borgida] and the Rome branch of LOA.

The General Formal Ontology (GFO), developed by Heinrich Herre and his colleagues of the research group Onto-Med in Leipzig, is a realistic ontology integrating processes and objects. It attempts to

include many aspects of recent philosophy, which is reflected both in its taxonomic tree and its axiomatizations. GFO allows for different axiomatizations of its categories (such as the existence of atomic time-intervals vs. dense time). The basic principles of GFO are published in the Onto-Med Report Nr. 8. Two GFO specialties, among others, are its account of persistence and its time model. Regarding persistence, the distinction between endurants (objects) and perdurants (processes) is made explicit within GFO by the introduction of a special category, a persistant. A persistant is a special category with the intention that its instances "remain identical" (over time). With respect to time, time intervals are taken as primitive in GFO, and time-points (called "time boundaries") as derived. Moreover, time-points may coincide, which is convenient for modeling instantaneous changes.

WordNet, a freely available database originally designed as a semantic network based on psycholinguistic principles, was expanded by addition of definitions and is now also viewed as a dictionary. It qualifies as an upper ontology by including the most general concepts as well as more specialized concepts, related to each other not only by the subsumption relations, but by other semantic relations as well, such as part-of and cause. However, unlike Cyc, it has not been formally axiomatized so as to make the logical relations between the concepts precise. It has been widely used in Natural Language Processing research.

The Suggested Upper Merged Ontology (SUMO) is another comprehensive ontology project. It includes an upper ontology, created by the IEEE working group P1600.1 (predominantly by Ian Niles and Adam Pease). It is extended with many domain ontologies and a complete set of links to WordNet. It is freely available.

Most of the existing ontologies have an upper level, above space/time horizon, and a specific level handling time and space. This two level approach asks for a very strong and consequent modelling approach. Best practice methodologies in the field are quite immature and have low uptake rate, this includes e.g. ISO 15926 [ISO 15926], ISO 10303 STEP [ISO 10303], UML as used in the Norwegian SERES project[18].

Further literature:

- Standard Upper Ontology Working Group (SUO WG), 4D Ontology. http://suo.ieee.org/SUO/SUO-4D/index.html
- Origins of The IEEE Standard Upper Ontology. http://www.ontologyportal.org/pubs/IJCAI2001.pdf

The Ontology of Spacetime (Philosophy and Foundations of Physics, Volume 1). This book contains selected papers from the First International Conference on the Ontology of Spacetime. http://www.spacetimesociety.org/

18.4 Ontology based data integration

An important aspect of the interoperability problem is that of integrating data from multiple heterogeneous sources. As interoperability is essentially about how to make different systems work together by sharing data and/or processes over a protocol, interoperability depends crucially on a common understanding, across the systems involved, of the *meaning* of the shareable resources. An

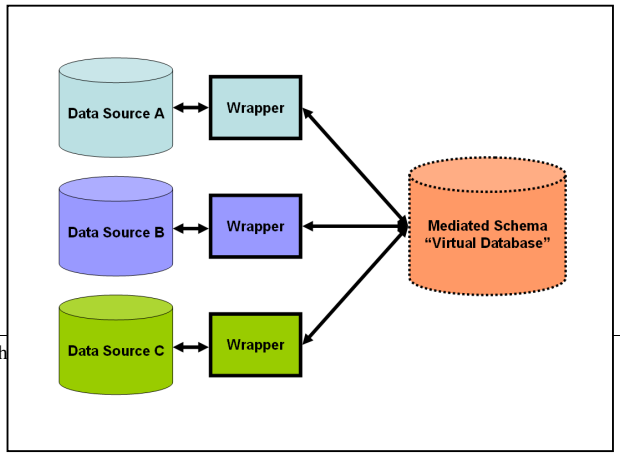
[18] Semantic register for electronic services, hosted at The Brønnøysund Register Center, http://www.brreg.no/samordning/semantikk/

inspection of different data repositories frequently reveals a plethora of structural and semantical interconnections. However, it is rarely the case that the formats involved are completely aligned, in the sense that it is possible for one system to understand and utilize, without further ado, the information contained in another.

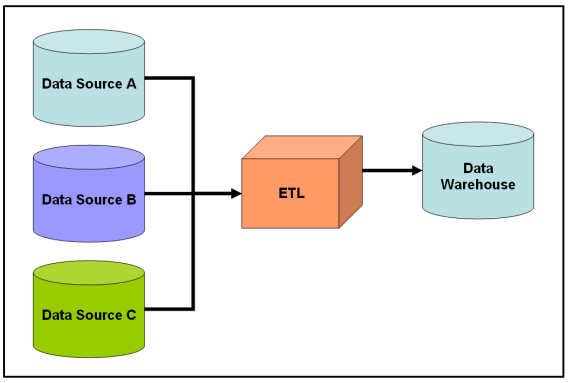
The choice of a format may be influenced by many factors. For instance, the uses to which the data is meant to be put may influence the granularity level of the basic units of data, as well as the overall semantics of the data structure. As a simple example consider an ordinary Web-page. A Web-page *does* have a semantics. The way the page is rendered by a browser is a result of the markup of its content in terms of the interrelations between the basic constituents of a page. However, a Web-page clearly has only a *presentation oriented* semantics. It contains no structure if it does not contribute directly to the description of the layout of the page. Thus values such as dates, zip codes or ISBN numbers, are not annotated as such, and are therefore semantically indistinguishable. For a computer system to utilise, say, geographical or historical information contained in a simple Web-page, therefore, a machine-readable representation has to be built on top of it that explains the meaning of the relevant data items.

In general terms, the problem of data integration is to design an architecture capable of bringing together information sources that may contain the data needed for a given task (or set of tasks). Once a set of sources have been found that bear importantly on the problem at hand, access to the data therein has to be provided. This means that each of the information sources will have to work together, to satisfy the requests of the system that is querying the information. In abstract terms, therefore, a data integration system is a virtual collocation of heterogeneous but related types of information resources, designed to present a unified view of physically and/or logically distinct data repositories. The idea is to free the user from tasks such as finding the relevant data sources, interacting with each source in isolation, and selecting, cleaning and combining data from multiple sources.

Formally a data integration system is a triple where G is the the target schema (henceforth called the *global* schema), S is a heterogeneous set of source schemas, and M is the mapping that correlates a query over the source schema with a query over the global schema. When users pose queries over the data integration system, therefore, they pose queries over G and the mapping then delegates the query to a source. It should be noted that the various data sources will usually have to be wrapped in mediator code. For instance a simple Web-page does not have a semantically significant structure at all (not counting visual rendering), so if it is to be used as a data source in an integrated system, an adapter has to be written that extracts and reformats the information in a way that makes it accessible it through the global interface.



The cornerstone of all data integration systems, as defined above, is thus the mapping *M* that correlates a query over a wrapper with a query over the global schema (called *mediates* schema in the illustartion above). Depending on the direction of this mapping, one can distinguish between two main approaches to data integration, namely the *Global as View* or GAV approach and the *Local as View* or LAV approach. They are called 'as view', since they both consider the data sources, suitably wrapped, as a set of views in the global schema. This contrasts with more traditional approaches such as *data warehosing*, where data from several sources are extracted transformed an loaded into a central storage that can be queried through a single schema:



Now, on the GAV approach, the global schema is modeled as a structure-preserving abstraction over the different sources in *S*. That is, the global schema expresses the logical relations between the data in the sources based on the structure of the wrapping of that source. In other words, the mapping *M* associates to each query over *G* a query over a source schema. Query processing becomes a straightforward operation because the associations between *G* and *S* are well-defined. The burden of complexity is placed on implementing mediator code instructing the data integration system exactly how to retrieve elements from the source databases. If any new sources are added to the system, considerable effort may be necessary to update the mediator, thus the GAV approach should be favored in cases where the sources are not likely to change. On a LAV approach, on the other hand, the direction of the mapping is reversed. Instead of mapping queries over a global schema to queries

over a source schema, a LAV system models a source database in terms of a set of queries over *G*. I. e. the direction of the mapping is reversed. Hence, *M* associates to each element in a source database a query over *G*. Another way to put it is to say that each of the sources is responsible for setting up its own mediator that allows it to be accessed via the global schema. LAV modelling is therefore more modular, and new sources can be added with far less work than in a GAV system. On the negative side, the mapping between the sources and the global schema is no longer well-defined, which will hamper a theoretical study of the approach.

Now, an *ontology based* data integration system, is a system that facilitates data integration and interoperability by encoding the semantic properties of the various types of data and their logical interrelations. According to the Web Services Modelling Framework (WSMF), interoperability problems can be grouped into three levels: data, process and protocol. The first of these in turn subdivides into the following categories:

- Format conversion, which is the problem of how to subsume different languages such as plain XML, RDF/OWL and relational databases schemas under a single formalism.
- Schema mapping, which is the problem of correlating synonymous units in different representations.
- Identity resolution, which is the problem of resolving references in such a way that homonyms are not conflated and synonyms are identified.

Ontologies can be used to provide a solution to all three of them, and is very often the most naturalchoice. An expressive ontology can be used as a kind of Esperanto for the schemata and data in the scope of the integration mechanism. As an example, consider a case of value conversion, where we want "28.11.08" to match "28th of November 2008". The matching is a non-trivial operation, in so far as the system needs to know that 11 maps to November and 08 to 2008 etc. Now, although this can be achieved with simple 'screenscraping' algorithms, a more principled, scalable, and extensible approach is to subsume each of the signifcant parts of a date object under a concept that describes how it is to be interpreted, as well as how it is related to other objects. As a result, date objects will have the same *logic* independently of their particular representation. Summing up so far, therefore, an ontology based data integration system uses an ontology for the explicit description of the information source semantics, thereby facilitating interoperability through a common understanding, on behalf of the systems involved, of the meaning of the shareable resources.

Needless to say, the problem as such of integrating data is not essentially tied to ontologies, and several other approaches are currently explored. A notable example is Clio, a joint research project at the University of Toronto and IBM's Almaden Research Center lead by Ronald Fagin and colleagues. A principal outcome of the Clio project is a software tool with a graphical user interface that creates mappings between two data representations semi-automatically (i.e., with user input). More specifically, the Clio *application* takes as input a target schema and a source schema, and generates queries based on correspondences perceived and input by the user. These queries take data from the source and transform it to match the target schema, cleansing and transforming it as needed to be compatible with existing data visible through that schema. In other words, the correspondences input by the user enables the Clio system to map a query in the source schema to one in the target [Hernández]. As a result the target schema provides a view over a dataset that comprises (at least some of) the data instantiating the source schema. Here the role of the mediator is played by the human user rather than by a semantic model.

The recent trend in data integration is to use description logics as ontology languages (this trend is currently spearheaded by the TONES project and its leading researchers (e.g. Diego Calvanese and Maurizio Lenzerini)). Description logics are very well suited for data modelling at the conceptual level, and have sufficient expressive power to model the structure of most of the common data-storage formats, such as for instance tables in a relational database. It is therefore a natural candidate for the specification of *inter-schema relationships,* that is, for the specification of commonalities and logical relationships between the structure of different databases. One can, for instance, use subsumption axioms on indexed relations to express such connections. More specifically, let R_i and S_j be two relations, indexed by their respective source databases i and j. If the designer of the integrated system knows that the attributes of R is a subset of the attributes of S, he can include an axiom

$$R_i \leq S_j$$

The axiom states that every object in i that satisfies R satisfies S in j provided it exists in j. The computational tractability of such DL schemas allows sophisticated forms of reasoning on inter-schema relationships, e.g. for inferring those extensional relationships between concepts that are implied by conceptual interdependencies. The task of querying a global schema over a set of source databases, therefore, comes down resolving a concept description and retrieving the corresponding records from the source databases [Borgida].

18.5 Referents and identity criteria

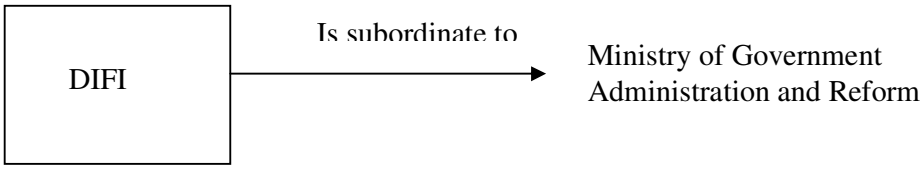
Vocabularies or terminologies are the set of properties, or meta-data elements defined by a group or an information community. At the heart of the problem of semantic interoperability is the question of how to equip each element in a vocabulary with a unique identity. Indeed, the question of the identity of a subject is at the heart of all knowledge interchange, be it between humans or computers. In order to communicate internally, a community agrees (to a reasonable extent) on a set of terms and their meanings. Although it is usually clear to the members what the elements in their shared vocabulary mean, the ability to standardise terminology becomes urgent once a vocabulary is reused and shared among disparate information communities wishing to use it for different ends. Ontologies will not be reusable unless they are based on a reliable and unambiguous identification mechanism for the things about which they speak. The same applies to thesauri, registries, catalogues, and directories. Software applications (for instance software agents) that collect or aggregate information and knowledge will not scale beyond a closely controlled environment unless the identification problem is solved.

18.5.1 URI-based standards

To meet the identity challenge most Web-oriented representation languages are based on URIs (*Uniform Resource Identifiers*). This is true of the W3C recommendations RDF (*Resource Description Framework*) and OWL (*Web Ontology Language*) as well as the ISO standard XTM (*XML Topic Maps*). A uniform resource identifier is a compact string of characters which, by design, identifies one unique resource on the Internet. URIs are a cornerstone of Web architecture, providing identification that is common across the Web. A main goal of the Web, since its inception, has been to build a global community in which any party can share information with any other party, and for that to become a reality a single global identification system is needed.

The above mentioned representation languages bind a subject to a URI in slightly different ways:

- The Resource Description Framework, as the name indicates, is intended as a model for describing internet resources, that is, it is intended as a standard for annotation of Web-accessible data. The basic RDF model is very simple. RDF defines a *resource* as any object that is uniquely identifiable by a URI. The properties associated with resources are identified by *property types*, and property types have corresponding *values*. The result is a so-called RDF triple consisting of a subject, a predicate and an object where the subject uses a URI as a binding point for its identity. A typical RDF triple, represented graphically, would look something like this



Here, the box is the subject, in this case the Agency for Public Management and eGovernment (aka. DIFI), the arrow represents the property of being subordinate to, and the distal end of the arrow is the object, that is, the superordinate Ministry of Government Administration and Reform. The RDF standard requires all boxes, that is all declared resources to be uniquely identified by a URI, in effect turning triples into modular units that can be exchanged and composed.

RDF uses the syntax rdf: about to point to a resource or subject. In the example above, for instance one would declare a *description* to be *about* DIFI in the following way:

<RDF: Description RDF: about=http://www.difi.no/difi.aspx?m=42591/>

where the value of the about-property is the address of the DIFI portal. The same basic mechanism is adopted by the web ontology language OWL, which has RDF as a language fragment. Note, that the use of the rdf:about syntax is a critical element in the creation of a distributed ontology; one ontology may refer to elements defined in a different ontology, thereby creating points of intersection over which information may be integrated an shared.

Although the use of unique resource identifiers is a simple and compelling strategy, it is not entirely free of problems. In order to be useful an ontology needs to refer to other things besides digital resources. An ontology usually models a domain, say privacy law, that to a large extent overlaps the 'analogue' world. It may therefore also need to refer to things, institutions and people. Alas, this makes identity criteria ambigous because there is no way to tell whether a URI is intended to refer to, say, an actual web page, or to that which the web page is about. Say for instance, that one wishes to refer to the Norwegian Prime Minister by binding a corresponding element in the ontology to an article on Wikipedia. Then how do we know that the URI is not meant to reference the *article* itself? The XTM

standard attempts to face up to this challenge by distinguishing between *subject indicators* and *resource references*:

- In Topic Maps, the central organizing unit it called a *topic*, which is roughly equivalent to a *subject* in RDF. A topic has a name, indicated by an *id* attribute, and it may be bound to a referent in either of two ways. Returning to the preceding example, say it is the Prime Minister himself you want to refer to. Then a declaration such as the following should be used:

```
<topic id=”jens_stoltenberg”>
   ….
   <subjectIndicatorRef  xlink: href=” http://no.wikipedia.org/wiki/Jens_Stoltenberg\>
   ….
</topic>
```

If, on the other hand, it is the article itself you have in mind, XTM offers the following variant

```
<topic id=”jens_stoltenberg_article”>
   ….
   <resourceRef  xlink: href=” http://no.wikipedia.org/wiki/Jens_Stoltenberg\>
   ….
</topic>
```

18.5.2 Published subject indicators

In general, a reference may be inteded to point to a variety of things; a resource, an object and a concept. The problem is that all of these must be identified by a URI and therefore by a digital *representation* of the thing. Now, URIs *are* unique, so however the reference is meant to be interpreted, it *has* been fixed once a URI is assigned. Hence the problem of disambiguating identity criteria boils down to that of standardising the interpretation of references. A general solution was suggested in the XTM 1.0 specification, in the form of sets of *published subject indicators.* Briefly put, a published subject is a any subject for which an adressable subject indicator has been made available for public use via a URI. A PSI is therefore any resource that has been published in order to provide a positive, unambigous indication of the identity of a subject for the purposes of facilitating topic map interchange and mergeability.

It should be clear that this strategy is not essentially tied to the topic map idiom, but could and should be considered as a general solution cutting across the various existing representation and markup languages out there. PSIs are basically just distinguished sets of URIs that all URI-based standards would be able to share. Hence any ontology in which identity is thus assigned could refer to a common standardised set of such, and use them to specify what an ontology is about. The identity of an element in an ontology would then be supported by names only in a relative form, valid only within a particular context or within a particular information community, whereas the absolute identity would be specified by referring to the correct PSI. Maintaining a public repository of PSIs is moreover an extensible approach that can be used for a variety of purposes:

- The principal application of PSIs would be to facilitate knowledge interchange, by specifiying the precise semantics of entities and relationships in an ontology. The points of intersections between different ontologies would always be precisely determinable, if at all, in effect enabling technology such as ontology alignment, merging and mapping.
- PSI repositories could also contain more than simple definitions of terms. One idea that suggests itself quite naturally is to organise a PSI repository as an ontology itself, proposing PSIs also for relations, properties and values. One could also imagine that a PSI repository would list its public users, and authorise different communities to access already existing ontologies as well as participating in constructing new ones. This could in turn be used to fuel the authority- and trust-building process that the semantic web relies on.

18.5.3 Life cycle standards

Examples of standards used for modelling information about referents during a life cycle are PLCS [ISO 10303] and ISO 15926 [ISO 15926]. The history of the standards goes 15-20 years back. The standards are in some use in special domains.

Examples of current research in the field are reflected in [Barry Smith-2] "Strategies for Referent Tracking in Electronic Health Records". The article elaborates the challenges of keeping all patient information in such a formal shape that it is suitable for search, statistic production of statistics and reasoning.

18.5.4 ID of referents

Semantic Technologies like Topic Map [Topic Maps] and RDF [RDF] (Resource Description Framework) use URIs (Uniform Resource Identifiers) to manage the difficult task of identifying both references and referents in a distributed environment.

Seen from the referent corner of the triangle an illustrated example of the problem is how to make sure that a record about a referent is unambiguously identifying the correct referent.

In a record many symbols are used, and many pointers to references are made. So if the main purpose of a record is to describe something relevant for a patient, many other references are made. Let us look into an example in an electronic patient record system

1. Record 1: The patient has a tumour type AAA in his left kidney, at time t0
2. Record 2: Kidney removed by surgery of type x, at time t1, by Dr. Terje Grimstad. Medicine of type A and B shall be used.
3. Record 3: Kidney transplantation at time t2. Medicine of type A and B shall be used.
4. Record 4: Transplantation failed after two months.
5. Record 5: Kidney transplantation, t3. Medicine of type A and C shall be used.

Among others the symbols "tumour", "kidney" and the identification "left" are used. For each record there will be a linkage between the record and the patient.

All the records are linked to the patient but not properly linked to the doctor or type of medicine used and only in one of the records "left kidney" is used. Are we sure that all the records are related to left kidney? Most records have a timestamp, but also periods are used like "after two months". Further

details in the use of medicines are listed in other documentation, and a set of answers from laboratories are separate records or attached to a record.

Suggestions for how to improve the situation is made in articles like [Barry Smith-2] "Strategies for Referent Tracking in Electronic Health Records".

18.5.5 Ontology, identification of references

The ability to use unique identifiers when referring to a reference is a prerequisite for using ontologies in runtime environments. Further the IDs are needed in tasks of ontology alignment, merging, mapping etc. Use of OID (object identifier) or URI (Unique resource identifier) are commonly used in best practice.

18.5.6 Best practice

Brønnøysund Register Center uses global unique identifiers for companies as referents (Norwegian legal companies). The ID is called "organisasjonsnummer". But the challenges remains when other companies identify Norwegian companies by e.g. name alone. Another challenge is that a legal company may be located at several sites. An ID for each site or activity is established and is called "bedriftsnummer".

In healthcare sector, there has been established a "System of coding units in specialised health care[19]" (freely translated from Norwegian "Organisasjonskodeverk for spesialisthelsetjenesten") - OK2007, which defines a standardised way of uniquely identifying all organisational units in specialised health care using OID, in addition to a standardised way of characterising organisational units (theier "properties" in terms of e.g. what kind of specialties a unit has, what kind of health services a unit delivers etc.). Using this coding system, there has been established a "Register of units in specialised health care" (in Norwegian "Register over enheter i spesialisthelsetjenesten") – RESH. OK2007 and RESH are described in more details at http://www.shdir.no/norsk_pasientregister/resh/ (only in Norwegian).

18.6 Worldview and context

Different types of worldviews at the same point in time are for example based on:

- Purpose of a concept (its function). The purpose of a referent may differ depending on the context and the process it is a part of. E.g. a heavy lexicon book may be used for pressing leaves in a herbarium, but the lexicon is not originally made for that purpose. A computer is a kind of multi-tool, even an effective criminal tool for Internet based crime.
- Where a referent is (in space). E.g. in international food transport the location of the food has an influence on what information is needed to handle it properly and to avoid damage. This is because differences in temperature, humidity, bugs, salt water exposure during transportation affect the food quality. This means that handling the same type of food in Egypt and at Svalbard could be quite different.
- Cultural and religious context. E.g. whether you eat pork meat or not is dependent on different religious traditions and how the pig is slaughtered.

[19] It is difficult to find an "accurate" translation of the Norwegian word "spesialisthelsetjeneste".

- Legal context. How one type of food is made and dealt with differs in different markets based on legislations, e.g. what drugs or chemicals one may use differs based on what legislation one must follow.

Countermeasures are used to reduce the challenge of semantic interoperability related to collaborating actors having the different types of worldviews listed above. The countermeasures in it self changes. The two main types of countermeasures are (i) procedural and administrative countermeasures, and (ii) technology based countermeasures.

18.6.1 Worldview

> *The semantic web communities focus on what a thing is (its substance), and do not allow for a gradual historical shift from substance to function. As a result the AI and semantic web communities create data structures that assume a single world-view. Every thing is presented as if this is the way "it is" ontologically, rather than providing frameworks whereby what a thing "is", what it means, and how it relates to other things, change as the framework changes. This dimension is needed a) to explore the interplay between facts and the frameworks or world-views used to explain them and b) to explain a historical shift from a quest for a single ontology to a need for multiple ontologies. Needed is an approach where entities can evolve in meaning. [*Veltman*]*

The triangle of reference [Ogden] has evolved, and different alternative views of the triangle have been made [Veltman, p 17]. However, even if we define our concepts/references separated from the terms/symbols and referents, we are influenced by our worldview and our purpose of modelling.

If a border of a country is disputed, then there will be different worldviews of the border. The same is the case in a legal dispute, e.g. if a certain delivery meets the completion criteria in a contract or not. The two parties may have different views of the concept of what should be delivered according to an agreement. This leads us to the need for managing multiple concepts in ontologies, and to the need for a mechanism for managing ontology merging and comparing.

The author of a record had an intention writing it. The reader's knowledge of the writer's intention will influence which worldview he will try to use as basis for his understanding the content.

Another example could be found in the feminine literary history. The knowledge base before feminine literature became available was heavily influenced by men's view of the world. We now lack the feminine worldview in a large part of our cultural history; we more or less have the history and worldview of men [Møller].

A systematic approach for managing ontologies with different worldviews covering the same referents would be of great interest to solve semantic interoperability. The methods found so far on ontology merge and mapping may be used for handling several ontologies covering the same referents and with different worldviews.

18.6.2 Context information

Handling context of a referent is related to whether the referent is material or not. *A material object* has a physical life-cycle, and information related to the referent will usually relate to the referent in a certain state or usage in a certain phase of its lifecycle. Examples of material referents could be a ship, a book or a person. A material object can not be fully digitally represented, but information about the object may be digitalized. A digital copy may be transformed to a material object, e.g. the digital original of a book can be printed and become a physical copy of the book. The whole content of a book could be digitally represented, but the physical book will still be physical and have some properties different from the digital copy.

An immaterial referent could be a house insurance, email, bank transaction or a video stream. The house is material, the insurance contract can be transformed to a material paper record, but the house insurance in it self is not physical[20]. An email can be printed, and than the printed representation of the email is physical. Characteristics of the *immaterial* objects are that they do not have any physical representations in real life that the record can relate to.

18.6.3 Semantic dimensions

Inspired from the field of philosophy the six basic questions – Why? How? When? Where? Who? What? – could have been used to structure this report. Some literature does focus on this problem breakdown structure.

> *Libraries typically offer access via author catalogues (Who?) and title catalogues (What?). In addition, libraries such as the Herzog August Bibliothek in Wolfenbüttel, offer access chronologically (When?) and via locations of publication (Where?). Search engines such as Artefacts Canada have begun to use such questions for searching. [*Veltman*]*

Systematic search using the six basic questions and their variants would greatly expand the scope and the precision of searching and could be of great help as countermeasure for preserving semantic value.

19 DATA AND SEMANTIC MODELS

19.1 Records & Data

A record is text, pictures, numbers etc. describing some aspects of a referent(s). In this chapter we present some issues related to a record itself and how it is influenced by time.

19.1.1 Data quality

Research within data quality and completeness of information operates with e.g. levels of what records conform to, and lists a set of criteria to be fulfilled in order to become conform. ISO 2382-8 defines data quality as follows:

> *Data quality is "the correctness, timeliness, accuracy, completeness, relevance, and accessibility that make data appropriate for their use"*

Different levels of data quality are categorized in [Price] as:

[20] This report does not go in depth in discussing what material is, or what it is not. We believe the answer depend on worldview.

> *"the three semiotic levels—syntactic, semantic, and pragmatic—describing respectively (1) form, (2) meaning, and (3) application (i.e. use or interpretation) of a sign can be used to define corresponding quality categories based respectively on (1) conformance to database rules, (2) correspondence to external (e.g. real-world) phenomena, and (3) suitability for use."*

In [Yang] "Journey to Data Quality" methodologies and quality metrics based on research are presented. This is presented along with case studies and suggestions on how to close the gap between the current quality level and the level you wish. An interesting overview of "Ten root conditions of data Quality" is elaborated, and consequences are discussed. The authors suggest treating information about a product as a product itself, not a bi-product, and present methodology and notations to make information product maps. The technique is used on example cases.

A relevant standard initiative is IS0 8000 Data Quality [ISO 8000].

A reflection: Every record is a referent, and also every ontology is a referent.

19.1.2 Data and model transformation issues

Once material is digital, it can be translated to other formats, representations and media. This is called transformation of a record. But also models and ontologies can be transformed from being built according to rules of one metamodel and transformed to be according to another metamodel (see e.g. OMG MOF (OMG MOF)). This holds true for many cases, but e.g. semantics contained in a written speech in a play, will not capture the speaker's way of performing the speech. A video of a play will contain more context information than an audio recording of the same play, but the video does not contain the theatre atmosphere and building architecture, audience, smell etc. Much of the semantic is in the way a speech is performed, the context the speech is held in, gestures of the speaker etc. This means that a digital text record of a speech contains less or at least other semantics than an audio file of the same speech, and also less semantics than a video of the play.

On the other hand, the audio part of a video could be used as a complete audio file.

A digital report can become a printed paper copy, with (under some circumstances) identical attributes to other paper copies.

Based on the chapter above we believe transforming could change the semantics of a record. Since many preservation regimes have transformation as one of several mechanisms for managing records, the semantic preservation should not be neglected during transformation.

19.1.3 Metadata

19.1.3.1 What is metadata

From http://en.wikipedia.org/wiki/Metadata :

Metadata (meta data, or sometimes metainformation) is "data about data", of any sort in any media. An item of metadata may describe an individual datum, or content item, or a collection of data including multiple content items and hierarchical levels, for example a database schema. In data processing, metadata is definitional data that provides information about or documentation of other data managed within an application or environment.

For example, metadata would document data about data elements or attributes, (name, size, data type, etc) and data about records or data structures (length, fields, columns, etc) and data about data (where it is located, how it is associated, ownership, etc.). Metadata may include descriptive information about the context, quality and condition, or characteristics of the data.

The term was introduced intuitively, without a formal definition. Because of that, today there are various definitions. The most common one is the literal translation:

- "Data about data are referred to as metadata."[21]

Example: "12345" is data, and with no additional context is meaningless. When "12345" is given a meaningful name (metadata) of "ZIP code", one can understand (at least in the United States, and further placing "ZIP code" within the context of a postal address) that "12345" refers to the General Electric plant in Schenectady, New York.

As for most people the difference between data and information is merely a philosophical one of no relevance in practical use, other definitions are:

- Metadata is information about data.
- Metadata is information about information.
- Metadata contains information about that data or other data.

The metadata concept has been extended into the world of systems to include any "data about data": the names of tables, columns, programs, and the like. Different views of this "system metadata" are detailed below, but beyond that is the recognition that metadata can describe all aspects of systems: data, activities, people and organizations involved, locations of data and processes, access methods, limitations, timing and events, as well as motivation and rules.

Fundamentally, then, metadata is "the data that describe the structure and workings of an organization's use of information, and which describe the systems it uses to manage that information". To do a model of metadata is to do an "Enterprise model" of the information technology industry itself.[22]

19.1.3.2 Types of metadata

Metadata can be classified by:[23]

- Content. Metadata can either describe the *resource* itself (for example, name and size of a file) or the *content* of the resource (for example, "This video shows a boy playing football").
- Mutability. With respect to the whole resource, metadata can be either *immutable* (for example, the "Title" of a video does not change as the video itself is being played) or *mutable* (the "Scene description" does change).
- Logical function. There are three layers of logical function: at the bottom the *subsymbolic* layer that contains the raw data itself, then the *symbolic* layer with metadata describing the raw data, and on the top the *logical* layer containing metadata that allows logical reasoning using the symbolic layer.

[21] James Martin, Strategic Data Planning Methodologies, Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1982, p.127

[22] William R. Durrell, Data Administration: A Practical Guide to Data Administration, McGraw-Hill, 1985

[23] http://en.wikipedia.org/wiki/Metadata

Metadata can in principle be any data of any relevance for any handling of our data elements. Supplementary types could be metadata for definition, documentation, technical system aspects, usage, administration, identification, security, presentation, references to ontologies, various contexts, information governance regime, purpose, legal aspects, intention of writer, etc.

19.1.3.3 Use of metadata

Each time a data element is used, sufficient metadata have to be available. The amount of metadata needed depends on each particular usage of the data element. E.g. different metadata might be needed if the data element is part of a calculation compared to the need if the data element is displayed on a screen.

In the context of interoperability with and within the public sector, metadata are used in a variety of situations:

- Forms reporting of information to government agencies (e.g. Altinn).
    - Handle the reception of the forms. Metadata (e.g. XML schemas) are extensively used to produce the application to control reception of the forms.
    - Fill information into the fields of a form. The user needs metadata (e.g. form user guidelines) in order to understand what to report.
- Databases.
  Metadata (e.g. database schemas) are used to facilitate the use and management of the database.
- IT applications.
  Metadata (e.g. java classes) are used to control the data processing.
- Collaboration of applications.
  Metadata (e.g. webService description) are used to control the interchange of information.
- Data quality.
  Both for forms, databases, and applications metadata (e.g. data types, constraints) can be used to detect incorrect or inconsistent data.
- Enterprise modeling.
  Good documented metadata (often at a non-technical abstraction level) is an essential part of obtaining an overall understanding of the information processing of the enterprise. Such understanding is important in the strategic planning and development of the enterprise, including improved utilization of data, collaboration with external sources, improved ability to change and modify, and preservation of personnel knowledge.

Since the need for metadata depends on the usage of the data, the form of the metadata will normally be different for the various situations above. However, it is important to obtain an overall consistent set of metadata.

19.1.3.4 Representation of metadata

Metadata exist and are represented in many ways: data type, tag in an XML message, XML schema, database schema, Java class, variable identifier, ORid, implicit assumption, documentation, in the head of a user, position on a screen, values of other data, incorporated in program code, etc.

The representation of metadata is

- more or less explicit,
- more or less structured,
- More or less formal.

Implicit and badly documented metadata are like an old undocumented Cobol program developed by now deceased programmers.

An important question is: ”how suitable is the representation of the metadata?”

Implicit and poorly structured metadata

- are the core of problems within semantic interoperability,
- reduce the potential usability of the data,
- make analysis of organizational interoperability more difficult,
- increase the probability of misusing data,
- make evolution of existing IT systems more difficult.

Formal and well structured metadata support

- automatic generation of IT system components (schema, code, data check),
- systems analysis,
- understanding.

19.1.3.5 The importance of metadata

Metadata have to be available dependent on the use of the data element. Metadata also have to be available when developing IT components operating on the given type of data. Thus, metadata are inherently necessary in all information processing. The quality of the metadata of an enterprise is directly related to the quality of the information processing of the enterprise.

The importance of metadata:[24]

“Metadata help you find, access, understand, and utilize the data.

Metadata:

- make the data easier to manage
- make data more useful to more people
- promote human and machine understanding

Metadata provide data-related information in a form that can be read by a human, or used in automated processing. Exposing the data in a formal, standards-compliant way to other interested scientists will increase the possibilities for solving problems, and help move science forward.

Metadata is the key to ensuring that resources will survive and continue to be accessible into the future.

[24] http://marinemetadata.org/guides/mdataintro/mdataimportance

Metadata will help raise awareness of the quality data and activities in your organization. In turn, this could help establish new collaborations which further utilize your valuable data assets."

19.1.4 Best practice and standards

Relevant standards for best practice for metadata are

- ISO 11179 Information technology Metadata Registries [ISO 11179]
- Dublin core [Dublin Core]
- DDI, Data Documentation Initiative [DDI]. DDI is an effort to establish an international XML-based standard for the content, presentation, transport, and preservation of documentation for datasets in the social and behavioural sciences
- Semantic annotation WSDL by W3C [SAWSDL]
- Neuchâtel Terminology Model, Classification database object types and their attributes [Neuchâtel]

19.2 Information Governance

Data or records about a referent: Are text, pictures, numbers, videos, models etc. describing some aspects of a referent. The data/record itself contains symbols. E.g. a record related to the tree I have in my garden could describe what type of tree it is, when it was planted, who planted it, how it was planted, what is the need for yearly maintenance, how it should be treated to have a long life etc.

As defined by the UK NHS (National Health System) [NHS], Information Governance ensures necessary safeguards for, and appropriate use of, all information assents relevant for operational and financial valuation of an enterprise. This encompasses a number of issues tied to financial, cultural and legal aspects of organisational interoperability.

The current focus in the field of Information Governance seems divided between Data Governance, Information Security Governance, IT Governance and Internet Governance. In the following subsections we briefly describe each subfield. In the next chapter we outline a few examples of best practise activities.

Data Governance is a quality control discipline for assessing, managing, using, improving, monitoring, maintaining, and protecting organisational information[25] It is a system of decision rights and accountabilities for information-related processes, executed according to agreed-upon models which describe who can take what actions with what information, and when, under what circumstances, using what methods [DGI]:

Data governance encompasses the people, processes, and information technology required to create a consistent and proper handling of an organisation's data across the business enterprise, including the goals of:

- Increasing consistency and confidence in decision making
- Decreasing the risk of regulatory fines
- Improving data security
- Maximizing the income generation potential of data

[25] http://www-01.ibm.com/software/tivoli/governance/servicemanagement/data-governance.html

- Designating accountability for information quality

These goals are realized by the implementation of Data governance programs, or initiatives to improve data quality by assigning a team responsibility for data's accuracy, accessibility, consistency, and completeness, among other metrics. This team usually consists of executive leadership, project management, line-of-business managers, and data stewards. The team usually employs some form of methodology for tracking and improving enterprise data, such as Six Sigma, and tools for data mapping, profiling, cleansing, and monitoring data.

Data governance initiatives may be aimed at achieving a number of objectives including offering better visibility to internal and external customers (such as supply chain management), compliance with regulatory law, improving operations after rapid company growth or corporate mergers, or to aid the efficiency of enterprise knowledge workers by reducing confusion and error and increasing their scope of knowledge. Many data governance initiatives are also inspired by past attempts to fix information quality at the departmental level, leading to incongruent and redundant data quality processes. Most large companies have many applications and databases that can't easily share information. Therefore, knowledge workers within large organisations often don't have access to the information they need to best do their jobs. When they do have access to the data, the data quality may be poor. By setting up a data governance practice or Corporate Data Authority, these problems can be mitigated.

The structure of a data governance initiative will vary not only with the size of the organisation, but with the desired objectives or the 'focus areas' [DGI] of the effort.

An governance regime for collaborating organisations must also cover topics as:

- Information security
- Information Technology Governance
- Internet governance

Information security means protecting information and information systems from unauthorized access, use, disclosure, disruption, modification, or destruction [Hopwood]. The terms information security and computer security are frequently incorrectly used interchangeably. Information security is concerned with the confidentiality, integrity and availability of data regardless of the form the data may take: electronic, print, or other forms. Computer security can focus on ensuring the availability and correct operation of a computer system without concern for the information stored or processed by the computer.

Information Technology Governance, IT Governance or ICT (Information & Communications Technology) Governance, is a subset discipline of Corporate Governance focused on information technology (IT) systems and their performance and risk management. The rising interest in IT governance is partly due to compliance initiatives, for instance Sarbanes-Oxley in the USA and Basel II in Europe, as well as the acknowledgment that IT projects can easily get out of control and profoundly affect the performance of an organisation.

Internet Governance, the definition has been contested by differing groups across political and ideological lines. One of the key debates centers on the authority and participation of certain actors, such as national governments and corporate entities, to play a role in the Internet's governance.

A Working Group established after a United Nations-initiated World Summit on the Information Society (WSIS) proposed the following definition of Internet governance as part of its June 2005 report:

*Internet governance is the development and application by Governments, the private sector and civil society, in their respective roles, of shared principles, norms, rules, decision-making procedures, and programmes that shape the evolution and use of the Internet [*DataGov*]*

Law professor Yochai Benkler developed a framework for conceptualizing the idea of Internet governance through the idea of three "layers" of governance: the "physical infrastructure" layer through which information travels; the "code" or "logical" layer that controls the infrastructure; and the "content" layer, which contains the information that runs through the network [MDM conf].

20 BEST PRACTICE

A number of Norwegian public agencies have in different ways established routines and systems for managing their more or less formal semantic models, their information models, metadata, master data etc. Different case-reports from the Semicolon project describe the current situations at the different partners.

20.1 Best practice – healthcare sector

The Norwegian healthcare sector is a complicated and complex sector. It consists of many actors who are autonomous entities. The Norwegian health sector itself could actually be considered as a "miniature of the Norwegian public sector". The description in this section may therefore not necessarily represent a complete picture of what is the best practice in the Norwegian healthcare sector (if there ever will be a complete picture).

In national IT-strategies[26] for healthcare sector, a "comprehensive and well-defined information base" for the sector is defined as one of the major priority areas.

On behalf of the national healthcare authorities, KITH established and maintains a metadata/semantics repository for the Norwegian healthcare sector. This includes, among other things, concept definitions and classification systems / coding schemas. This work covers various healthcare professions/fields (medical, nursing etc.) as well as organizational/technical concepts and classifications. Though not explicitly stated, the "KITH-standards" on the structure (technical/syntactical/logical) and the content of the electronic messages are also considered as a part of the "comprehensive and well-defined information base" which is considered necessary for achieving interoperability in the Norwegian healthcare sector.

The main regime for establishing and maintaining the metadata/semantics in healthcare sector is as following:

- The work of establishing or revising/updating the content of this unified/common information base, is conducted as an open standardisation process with involvement of representatives from the implied user groups. When doing this, one aims at using or harmonising with existing

[26] Among which, "Te@mwork 2007 - Electronic Cooperation in the Health and Social Sector - National Strategy 2004-2007 for Norway", http://shdir.no/publikasjoner/handlingsplaner/te_mwork_2007___electronic_interaction_in_the_health_and_social_sector_48934

international standards, especially European standards that the Norwegian public sectors are obligated to use.

- The results are (intentionally) published through the metadatabase www.volven.no which is conform to ISO standard IS11179, and/or at the website of KITH (www.kith.no/kodeverk), the electronic search tool (http://finnkode.kith.no) which is free of use, etc.

Some defined methods are used in defining/modelling/maintaining this information base, which though is out of the scope of this section of this document.

See also chapter on KITH and www.volven.no in Part I of this report.

20.2 Best practice for egov Interoperability

In the following review of best practice we outline selected interoperability frameworks that (1) are in widespread use, and (2) have significant focus on issues related to semantics. These include the EIF initiatives in the EU, national European initiatives, the Athena framework, and the US FEA initiative. A number of other initiatives deserve mentioning, including national and state eGov initiatives in Australia and the DODAF framework of the US Department of Defence.

In 2008 Semantic Interoperability Centre Europe was established. IDABC launched the Semantic Interoperability Centre Europe (SEMIC.EU), which lets eGovernment and other pan-European collaborations exchange their knowledge and visions. The Semantic Interoperability Centre Europe (SEMIC.EU) is an eGovernment service initiated by the European Commission and managed by the eGovernment Services (IDABC) Unit.

SEMIC.EU operates as a network comprised of eGovernment projects and communities and is completely open for participation by all stakeholders in the field of public services and eGovernment. It fosters the collaboration of EU Member States' initiatives in the field and complements them as a pan-European service: Semantic solutions provided by stakeholders are harmonized and pooled. The projects involved provide and reuse solutions for seamless data interchange which preserves the original meaning of the data. This collaborative approach is chosen to avoid redundancies in optimizing data exchange between different systems and across national and language boundaries.

The website www.semic.eu is built around a repository of real-life solutions dubbed 'interoperability assets' which are provided by projects and organizations in the sector. Within SEMIC.EU, the models, taxonomies, XML schemas and other solutions to enhance meaningful data exchange, are subject to a standardized and supervised clearing process. This involves peer review and incremental enhancement of the assets. However, they remain available at any stage during the process making it possible to reuse them in different projects, contexts or domains at any time. For discussions around the provided assets and to debate more general questions around issues of (semantic) interoperability, an open forum is an integral part of the portal. It is supervised by trained staff and serves as a source of input for project and platform improvement, based on the idea of exchange of solutions to the challenge of semantic interoperability, with a clear focus on eGovernment in Europe. As an interactive service its guiding principles are sharing and collaboration. The public portal www.semic.eu was launched in June 2008. It follows a clear open source policy as it is represented by the European Commission's Open Source portal www.osor.eu.

It is also an implementation-oriented preparatory measure of standardisation measures. Built upon its function as a facilitator of European harmonization, SEMIC.EU participates in international

standardisation organizations. In addition, it is among the service's primary goals to provide scientific as well as practical information in all matters related to semantic interoperability.

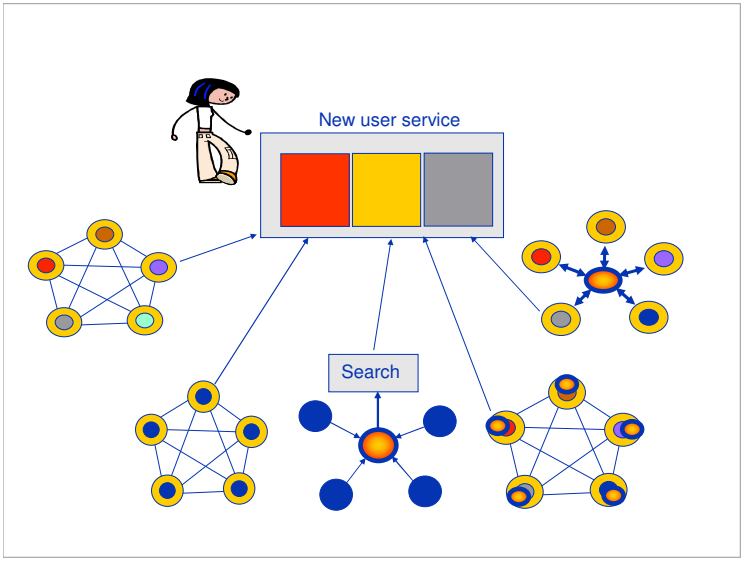
As a 'horizontal measure' of the IDABC program, SEMIC.EU is established as a permanent implementation of the principles stipulated in the 'European Interoperability Framework' (EIF).

20.3 Semantic interoperability architectures

Information models and ontologies can be implemented and maintained in many different ways in an enterprise. How use of these models is implemented and how groups of enterprises choose to maintain their private or common ontology impact the enterprises ability to collaborate.

> *Although service-oriented architectures go a long way toward providing interoperability in distributed, heterogeneous environments, managing semantic differences in such environments remains a challenge[Vetere].*

The figure below illustrates how different semantic and technical architectures implements semantic technologies and ontologies, and how this differences may lead to difficulties or incompatible collaboration. An enterprise will have to choose how to implement semantic technologies in his enterprise. Often his choice is influenced on what is common in his domain. E.g. when a laboratory in the health sector chooses his semantic technology approach his must comply to e.g. Norwegian health related specification to run his main business. When the enterprise then also needs to collaborate with banks, transportations services, public services and fulfil public obligations the pictures is often very complicated. You may mange to handle a few technical and semantically architectures, but the cost of handling many is still too high. Further research is needed even if there is much focus on how Service Oriented Architectures can be used to encapsulate and solve parts of this problem.



The figure illustrate how different domains choose to have peer to peer communication with or without a common ontology, whether the transformation services of data exchange is done locally or at a domain service, or whether the data exchange is done through a common hub which also does the

needed format and semantic transformations. And at the end, the user of the data has to compile the differences in semantics before he is able to use the data to perform his tasks.

20.4 Best practise on Information Governance

In this chapter we and outline a few examples of best practise activities. Needless to say, these are only a small sample of the many activities and initiatives that take place worldwide on different aspects of Information Governance, and many more references to ongoing work can be found from information published and updated on (an almost daily basis on) the Internet.

20.4.1 Implementation of Data Governance

Leaders of successful data governance programs declared in December 2006 at the Data Governance Conference in Orlando, Fl, that data governance is between 80 and 95 percent communication[27]. That stated, it is given that many of the objectives of a Data Governance program must be accomplished with appropriate tools. Many vendors are now positioning their products as Data Governance tools; due to the different focus areas of various data governance initiatives, any given tool may or may not be appropriate, in addition, many tools that are not marketed as governance tools address governance needs.

The IBM Data Governance Council [6]

> The IBM Data Governance Council is an organization formed by IBM consisting of companies, institutions and technology solution providers with the stated objective to build consistency and quality control in governance, which will help companies better protect critical data."

The Data Governance and Stewardship Community of Practice (DGS-COP)[7]

> The Data Governance and Stewardship Community of Practice is a vendor-neutral organization open to practitioners, stakeholders and academics, as well as vendors and consultants. The DGS-COP offers a large collection of data governance artifacts to members including case studies, metrics, dashboards, and maturity models as well as on-line events.

Data Governance Conferences[8]

> Two major conferences are held annually, the Data Governance Conference, held in 2008 in San Francisco, CA, USA, and the Data Governance Conference Europe, held in 2008 in London, England.

Master Data Management & Data Governance Conferences[9]

> Six major conferences are held annually, London, San Francisco, Sydney and Toronto in the spring, and Madrid, Frankfurt, and New York City in the fall. 2009 is the 4th annual iteration with more than 2,000 attendees per year receiving their data governance and master data management updates via this 2-3 day event.

Implementation of a Data Governance initiative may vary in scope as well as origin. Sometimes, an executive mandate will arise to initiate an enterprise wide effort, sometimes the mandate will be to create a pilot project or projects, limited in scope and objectives, aimed at either resolving existing issues or demonstrating value. Sometimes an initiative will originate lower down in the organisation's hierarchy, and will be deployed in a limited scope to demonstrate value to potential sponsors higher up in the organisation.

[27] http://en.wikipedia.org/wiki/Data_governance#Implementation

20.4.2 Implementation of Information Security Governance

ISO/IEC 27002 *Information technology - Security techniques - Code of practice for information security management* is an information security standard, and part of a growing family of ISO/IEC ISMS standards (the ISO/IEC 27000 series). It was published by the International Organisation for Standardization (ISO) and the International Electrotechnical Commission (IEC) as ISO/IEC 17799:2005 and subsequently renumbered ISO/IEC 27002:2005 in July 2007, bringing it into line with the other ISO/IEC 27000-series standards. The current standard is a revision of the version first published by ISO/IEC in 2000, which was a word-for-word copy of the British Standard (BS) 7799-1:1999.

ISO/IEC 27002 provides best practice recommendations on information security management for use by those who are responsible for initiating, implementing or maintaining Information Security Management Systems (ISMS). Information security is defined within the standard in the context of the C-I-A triad:

the preservation of confidentiality (ensuring that information is accessible only to those authorized to have access), integrity (safeguarding the accuracy and completeness of information and processing methods) and availability (ensuring that authorized users have access to information and associated assets when required).

ISO 17799:2005 standard is the most recently published revision of ISO's global security framework. This version significantly improves the already well-respected and comprehensive "Code of Practice for Information Security Management." It provides principles and guidelines for initiating, implementing, maintaining, and improving information security management throughout the enterprise. This includes best practices, control objectives and controls for a range of IT functions related to protecting information.

The ISO/IEC 27002:2005 Code of practice for information security management recommends the following be examined during a risk assessment:

- security policy,
- organisation of information security,
- asset management, human resources security,
- physical and environmental security,
- communications and operations management,
- access control,
- information systems acquisition,
- development and maintenance,
- information security incident management,
- business continuity management, and
- regulatory compliance.

With respect to identification and authentication for computer systems in use today, the Username is the most common form of identification and the Password is the most common form of authentication. Usernames and passwords have served their purpose but in our modern world they are no longer adequate. Usernames and passwords are slowly being replaced with more sophisticated authentication mechanisms.

After a person, program or computer has successfully been identified and authenticated then it must be determined what informational resources they are permitted to access and what actions they will be allowed to perform (run, view, create, delete, or change). Such *authorization* to access information and other computing services begins with administrative policies and procedures. The policies prescribe what information and computing services can be accessed, by whom, and under what conditions. The access control mechanisms are then configured to enforce these policies. Different computing systems are equipped with different kinds of access control mechanisms; some may offer a choice of different access control mechanisms. The access control mechanism a system offers will be based upon one of three approaches to access control or it may be derived from a combination of the three approaches.

The non-discretionary approach consolidates all access control under a centralized administration. The access to information and other resources is usually based on the individuals function (role) in the organisation or the tasks the individual must perform. The discretionary approach gives the creator or owner of the information resource the ability to control access to those resources. In the Mandatory access control approach, access is granted or denied bases upon the security classification assigned to the information resource.

Examples of common access control mechanisms in use today include Role-based access control available in many advanced Database Management Systems, simple file permissions provided in the UNIX and Windows operating systems, Group Policy Objects provided in Windows network systems, Kerberos, RADIUS, TACACS, and the simple access lists used in many firewalls and routers.

To be effective, policies and other security controls must be enforceable and upheld. Effective policies ensure that people are held **accountable** for their actions. All failed and successful authentication attempts must be logged, and all access to information must leave some type of audit trail.

In computing, e-Business and information security it is necessary to ensure that the data, transactions, communications or documents (electronic or physical) are genuine (i.e. they have not been forged or fabricated.). It is also important for authenticity to validate that both parties involved are who they claim they are.

In law, non-repudiation implies one's intention to fulfill their obligations to a contract. It also implies that one party of a transaction can not deny having received a transaction nor can the other party deny having sent a transaction.

Electronic commerce uses technology such as digital signatures and encryption to establish authenticity and non-repudiation. Digital signatures are also used in healthcare sector for communication containing sensitive health information.

20.4.3 Implementation of Information Technology Governance

After the widely reported collapse of Enron in 2000, and the alleged problems within Arthur Andersen and WorldCom, the duties and responsibilities of the boards of directors for public and privately held corporations were questioned. As a response to this, and to attempt to prevent similar problems from happening again, the US Sarbanes-Oxley Act (SOX) was written to stress the importance of business control and auditing. Sarbanes-Oxley and Basel-II in Europe have been catalysts for the development of the discipline of information technology governance since the early 2000s. However, the concerns

of Sarbanes Oxley (in particular Section 404) have less to do with IT decision rights as discussed by Weill and Ross [6] and more to do with operational control processes such as Change management.

What IT controls are most important for SOX compliance? A growing number of corporate IT organisations are finding at least some of the answers in recent iterations of two venerable standards frameworks: COBIT and ISO 17799:2005 (renamed to ISO 27005, as described in the previous section on Information Security implementation).

Control Objectives for Information and related Technologies (COBIT) is an open standard published by the IT Governance Institute and the Information Systems Audit and Control Association (ISACA).[15] A new version recently published, COBIT 4.0, emphasizes regulatory compliance as it relates to IT governance. ISACA, and describes COBIT as an IT governance framework with a supporting toolset that allows managers to bridge the gap between control requirements, technical issues and business risks.

COBIT provides a best practice framework for how to control, manage and measure 34 key IT practices. This framework includes high-level and detailed control objectives for each process, management guidelines (including process inputs and outputs, roles and responsibilities, and metrics), and process maturity models. A core emphasis of COBIT is aligning IT operations with strategic enterprise objectives and priorities to improve IT value delivery, resource management, business performance, efficiency and risk management.

The ISO 17799:2005 standard includes extensions that strengthen controls designed to protect the integrity of information from asset management and access control, to human resources security, security incident management and business continuity management. An important new requirement is an increased emphasis not only on the need to have good security controls, but also on the capability to validate the integrity of regulated information. It mandates validation through systematic auditing and monitoring of activity to prevent unauthorized access to sensitive corporate and customer information. Just as ISO 9000/9001 is used universally as a measure of production quality, ISO 17799:2005 is poised to play a similar role in the area of information integrity assurance.

Both COBIT and ISO 17799/2005 provide guidelines that are useful in helping companies determine how to think about the root requirements of compliance regulations and managing data risks. Developed specifically for IT organisations, these frameworks provide specific practices and guidelines for instituting controls aimed at ensuring the integrity of information assets.

Following Corporate Collapses in Australia around the same time, working groups were established to develop standards for Corporate Governance. A series of Australian Standards for Corporate Governance were published in 2003, these were:
Good Governance Principles (AS8000)
Fraud and Corruption Control (AS8001)
Organisational Codes of Conduct (AS8002)
Corporate Social Responsibility (AS8003)
Whistle Blower protection programs (AS8004)

AS8015 Corporate Governance of ICT, the Australian Standard for Corporate Governance of ICT, was published in January 2005. It was fast-track adopted as ISO/IEC 38500 in May 2008. AS8015 defines Corporate Governance of ICT as "The system by which the current and future use of ICT is directed and controlled. It involves evaluating and directing the plans for the use of ICT to support the organisation and monitoring this use to achieve plans. It includes the strategy and policies for using ICT within an organisation."

ISO/IEC 29382, Corporate Governance of Information and Communication Technology, was first published early in 2007 as a fast track candidate from the existing Australian standard AS8015. It was officially re-named ISO/IEC 38500 in April 2008. As is usual with international standards, it is intended to provide guiding principles to any organisation, regardless of size or sector.

ISO 38500 is now the international standard for the corporate governance of information technology has now been published. The original draft number for the standard of ISO 29382 has been discarded, and the official number of the new standard is ISO/IEC 38500. It draws upon a number of sources, chief of which is AS 8015:2005, which defines six principles (establish responsibilities, plan to best support the organisation, acquire validly, ensure performance when required, ensure conformance with rules, ensure respect for human factors).

ISO/IEC 38500:2008, corporate governance of information technology, is applicable to organisations of all sizes, including public and private companies, government entities, and not-for-profit organisations. This standard provides a framework for effective governance of IT to assist those at the highest level of organisations to understand and fulfil their legal, regulatory and ethical obligations in respect of their organisations' use of IT. The framework comprises definitions, principles and a model. It sets out six principles for good corporate governance of IT that express preferred behaviour to guide decision making:
(1) responsibility, (2) strategy, (3) acquisition, (4) performance, (5) conformance and (6) human behaviour.

The purpose of the standard is to promote effective, efficient and acceptable use of IT in all organisations by: assuring stakeholders that, if the standard is followed, they can have confidence in the organisation's corporate governance of IT informing and guiding directors in governing the use of IT in their organisation, and providing a basis for objective evaluation of the corporate governance of IT.

Practical application of these standards and principles typically takes place in IT Service Management (ITSM) programs, which is a discipline for managing information technology (IT) systems, philosophically centered on the *customer's perspective of IT's contribution to the business.* ITSM stands in deliberate contrast to technology-centered approaches to IT management and business interaction. ITSM is often equated with the Information Technology Infrastructure Library, (ITIL), an official publication of the Office of Government Commerce [16] in the United Kingdom. However, while a version of ITSM is a component of ITIL, ITIL also covers a number of related but distinct disciplines and the two are not synonymous.

ITIL is a set of concepts and policies for managing information technology (IT) infrastructure, development and operations. ITIL gives a detailed description of a number of important IT practices with comprehensive check lists, tasks and procedures that can be tailored to any IT organisation. The "Service Management" section of ITIL version 2 was made up of eleven different disciplines, split into two sections, Service Support and Service Delivery. This use of the term "Service Management" is how many in the world interpret ITSM, but again, there are other frameworks, and conversely, the entire ITIL library might be seen as IT Service Management in a larger sense. The new ITIL v3 rewrite has not similarly designated a subset as "Service Management."

20.4.4 Implementation of Internet Governance

To understand how the Internet is run today, it is necessary to know some of the key milestones of Internet governance.

The original ARPANET, one of the components which eventually evolved into the Internet, connected four Universities: University of California Los Angeles, University of California Santa Barbara , Stanford Research Institute and Utah University. The IMPs, interface minicomputers, were built in 1969 by Bolt, Beranek and Newman under a proposal by the US Department of Defense Advanced Research Projects Agency. By 1973 the ARPANET connected many more systems and included satellite links to Hawaii and Scandinavia, and a further link from Norway to London. It continued to grow in size, becoming more a utility than a research project. For this reason in 1975 it was transferred to the US Defense Communications Agency. During the development of ARPANET, a numbered series of Request for Comments (RFCs) memos documented technical decisions and methods of working as they evolved. The standards of today's Internet are still documented by RFCs, produced through the very process which evolved on ARPANET. The Internet protocol suite, developed between 1973 and 1977 with funding from ARPA, was intended to hide the differences between different underlying networks and allow many different applications to be used over the same network.

In 1979 the Internet Configuration Control Board was founded by DARPA to oversee the network's development. In 1984 it was renamed the Internet Advisory Board (IAB), and in 1986 it became the Internet Activities Board. RFC 801 describes how the US Department of Defense organized the replacement of ARPANET's Network Control Program by the new Internet Protocol in January 1983. In the same year, the military systems were removed to a distinct MILNET, and the Domain Name System was invented to manage the names and addresses of computers on the "ARPA Internet". The familiar top-level domains .gov, .mil, .edu, .org, .net, .com, and .int, and the two-letter country code top-level domains were deployed in 1984. Between 1984 and 1986 the US National Science Foundation created the NSFNET backbone, using TCP/IP, to connect their supercomputing centers. The combined network became widely known as the Internet.

Outside of the USA the dominant technology was X.25. The International Packet Switched Service, created in 1978, used X.25 and extended to Europe, Australia, Hong Kong, Canada, and the USA. It allowed individual users and companies to connect to a variety of mainframe systems, including Compuserve. Between 1979 and 1984, an approach known as Unix to Unix Copy Program grew to connect 940 hosts, using methods like X.25 links, ARPANET connections, and leased lines. Usenet News, a distributed discussion system, was a major use of UUCP. The Internet Engineering Task Force (IETF) was formed in 1986 by the US Government to develop and promote Internet standards. It initially consisted of researchers, but by the end of the year participation was open to anyone, and its business was largely carried on by email. By the end of 1989 Australia, Germany, Israel, Italy, Japan,

Mexico, the Netherlands, New Zealand, and the United Kingdom had connected to the Internet, which now contained over 160,000 hosts.

In 1990, ARPANET formally shut down, and in 1991 the NSF dropped its restrictions on commercial use of its part of the Internet. Commercial network providers began to interconnect, extending the Internet. In 1992 the Internet Society (ISOC) was founded, with a mission to *"assure the open development, evolution and use of the Internet for the benefit of all people throughout the world"*. Its members include individuals (anyone may join) as well as corporations, organisations, governments, and universities. The IAB was renamed the Internet *Architecture* Board, and became part of ISOC. The Internet Engineering Task Force also came under the ISOC umbrella. The IETF is currently overseen by the Internet Engineering Steering Group (IESG), and longer term research is carried on by the Internet Research Task Force and overseen by the Internet Research Steering Group.

Allocation of IP addresses was delegated to four Regional Internet Registries (RIRs):
American Registry for Internet Numbers (ARIN) for North America
Réseaux IP Européens - Network Coordination Centre (RIPE NCC) for Europe, the Middle East, and Central Asia Asia-Pacific Network Information Centre (APNIC) for Asia and the Pacific region Latin American and Caribbean Internet Addresses Registry (LACNIC) for Latin America and the Caribbean region In 1998, the IANA function was taken over by the Internet Corporation for Assigned Names and Numbers (ICANN), a newly created Californian non-profit corporation, set up in September 1998 by the US Government and awarded a contract by the US Department of Commerce. Initially two board members were elected by the Internet community at large, though this was changed by the rest of the board in 2002 in a thinly attended public meeting in Accra, in Ghana. In 2004 a new RIR, AfriNIC, was created to manage allocations for Africa. In 2002, a restructuring of the Internet Society gave more control to its corporate members.

At the first World Summit on the Information Society (WSIS) in Geneva 2003 the topic of Internet governance was put on the table. Since no general agreement existed even on the definition of what comprised Internet governance, United Nations Secretary General Kofi Annan set up a Working Group on Internet Governance (WGIG) to clarify the issues and report before the second part of the World Summit on the Information Society in Tunis 2005. After much controversial debate, participants agreed on a compromise to allow for wider international debate on the policy principles. They agreed to establish an Internet Governance Forum, to be convened by United Nations Secretary General before the end of the second quarter of the year 2006. The Greek government volunteered to host the first such meeting.

Today almost all Internet infrastructure is provided and owned by the private sector. Traffic is exchanged between these networks, at major interconnect points, in accordance with established Internet standards and commercial agreements.

21 SEMANTIC INTEROPERABILITY, MAIN METHODICAL AND TECHNOLOGICAL DIRECTIONS

The focus of Semicolon is semantic interoperability between collaborating organisations. We could expand the scope to cover e.g. data warehouse initiative, semantic search, semantic SOA etc, but in this chapter we have limited our scope to the core of the two directions:

- W3C initiative on semantic technologies and semantic related standards
- OMG Model driven architecture and use of Common Warehouse Metamodel

A short description of the initiatives is described below. There is ongoing research on how to combine the two initiatives.

21.1 Semantic technologies

Semantic technologies can be described as technologies that provide tools and methods to build more adaptive and flexible software by exploiting the meaning of the information at hand. One of the strongest drivers for semantic technologies at the moment is the work related to the Semantic Web. [FFI]

The core of the semantic technologies, the ontology, is a machine interpretable model(s) representing e.g. knowledge, behaviour and/ or meaning of things. These models may be engineered, manipulated, interpreted, etc in many ways giving us the opportunity to build systems and services on both data and models in new ways.

In the [FFI] report, five core semantic technologies are identified: Knowledge representation, ontologies, reasoning and rules, querying, and agents and services. Semantic technologies are expected to provide several interesting capabilities enhancing the potential of traditional information technologies regarding making more adaptive and flexible software.

Knowledge representation is in the context of semantic technologies usually identified with ontology construction, though the possibility of extending ontology languages with expressivity for the representation of other forms of knowledge is currently actively addressed by the research community, in particular related to rules, nonmonotonic inference, integrity constraints and epistemic concepts. The main function of reasoning is currently served by the tool support offered for ontology construction (see section 20.2), and to some extent on query answering. An application of ontology techniques for query answering that is of particular relevance for Semicolon is discussed above in section “Ontology based integration”. Use of software agents that operate on semantic specifications belong to the original vision of the Semantic web, but has so far seen very few adaptations.

The Semicolon project is in dialog with vendors like Cambridge Semantics and Zepheira on testing and using their product portfolios related to make demonstrators as part of our case work. We have also some knowledge on TopBraid composer and tools like Protegé.

In the report [David Provost, 2008], “A Global Review of the Industry and Leading Vendors”, a set of vendors and their offerings are listed, see table below.

Some Semantic technology vendors

Company	Solution	Middleware	Natural Lang Processing	Database	Platform	Ontology	Search	Consumer Web Service	Developer Web Service
Aduna	x	x		x					
The Calais Initiative			x						x
Cambridge Semantics	x	x							
Dow Jones Client Solutions	x					x	x		
Expert System			x						
Franz				x					
Mondeca					x				
Ontoprise	x	x	x	x	x	x	x		
Ontos			x		x	x			
OpenLink Software		x		x	x				
Primal Fusion								x	
Saltlux			x			x	x		
Sindice									x
Thetus	x				x				
TopQuadrant	x				x	x			
Twine/Radar Networks								x	
Yahoo!/SearchMonkey							x		x

Source: David Provost: A Global Review of the Industry and Leading Vendors, 2008

21.2 Model driven architecture and Common Warehouse Metamodel

Common Warehouse Metamodel (CWM) from OMG is a meta model that is intended for the modeling of common (for a set of applications e.g within an enterprise) meta data models and for the modeling interchange of meta data by means of transformations between models based on different meta models. Data warehouse is a repository of an organization's stored data. Data warehouses are designed to facilitate reporting and analysis. However, the means to retrieve and analyze data, to extract, transform and load data, and to manage the data dictionary are also considered essential components of a data warehousing system. Thus, data warehousing includes business intelligence tools, tools to extract, transform, and load data into the repository, and tools to manage and retrieve metadata.
The main usage scenario for CWM is to make centralized warehouse models that are common for a set of applicatios. As the centralized approach is abandoned in Semicolon, the interest in CWM is in its capabilities in defining different meta data models and transformations between these.

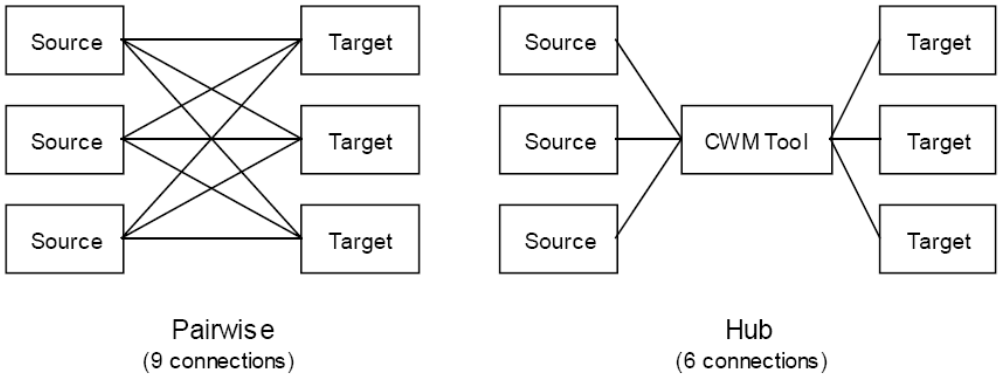


Figure 1. Data warehouse source and target arrangements.

Figure from CWM presentation by Doug Tolbert, Unisys Corporation.

As a meta model it defines a language at the same level as UML and OWL, and it is used for making models. It is based upon the foundation of UML, but it adds meta classes (and thereby language mechanisms) for modeling things like resources, relations, etc. It also includes meta classes for the modeling of transformations between models, and in this respect it is in line with the MDA approach of defining transformationsbetween models (based on different meta models) by means of models. CWM is based upon the MOF-related technologies of OMG and interchange is based upon XML/XMI.

Judging from (http://www.cwmforum.org/about.htm) it is uncertain how active the CWM community is; the most recent paper is from 2003, and the most recent event is from 2000.

If one is to consider the combined use of model based techniques (for constructive modeling) with semantic technology (for the intensional modeling), then ODM would be an alternative to CWM. It contains a standard meta-model for ontology modeling, a UML2 Profile for representing Ontologies, and mappings between ODM and the W3C OWL. However, although ODM is an OMG adopted technology, it is not supported by many tools. (http://wiki.eclipse.org/MDT/ODM-Proposal) describes a proposal for an Eclipse project on ODM.

22 CONCLUSIONS ON SEMANTIC TECHNOLOGIES

TBD in coming deliverables from the project.

23 REFERENCES

Alberts	Alberts, D.S. and Hayes, R.E. (2003). Power to the Edge, Command and Control in the Information Age. Information Age Transformation Series, CCRP Press
Athena DA4.2	Specification of interoperability framework and profiles, guidelines and best practices. Deliverable A4.2 from the Athena Project http://interop-vlab.eu/ei_public_deliverables/athena-deliverables/A4/d-a4.2/
Bachiechner 2006	Web Service Discovery – A Reality Check, DERI, Daniel Bachiechner, et al. 17. January 2006. http://www.sti-innsbruck.at/fileadmin/documents/DERI-TR-2006-01-17.pdf
Barry Smith	A Realism-Based Approach to the Evolution of Biomedical Ontologies. Werner Ceusters , Barry SMITH. forthcoming in Proc. AMIA Symp. 2006. http://ontology.buffalo.edu/bfo/Versioning.pdf
Barry Smith - 2	Strategies for Referent Tracking in Electronic Health Records. Werner Ceusters, Barry Smith. Preprint version of paper forthcoming in Journal of Biomedical Informatics
BEA 2008	The state of the BPM Market. BEA White paper.2008. www.bea.com/banners/solutions/bpm/L1-bp2.jsp
Bernstein 2008	Information integration in the Enterprice. Communication of the ACM September 2008. Philip A. Berstein and Laurea M. Haas.
Borgida	A. Borgida, M. Lenzerini and R. Rosati: “Description logics for databases”, in The Description Logic Handbook. Eds. F. Baader et al. Cambridge University Press 2004.
Brodie 2006	Semantically enabled Service-Oriented Architectures: A Manifesto and a paradigm

	shift in computer science. DERI Michael Brodie et. Al. 26. December 2006. www.deri.ie/fileadmin/documents/DERI-TR-2005-12-26.pdf
Christiansen	Christiansen, T.R. "Modeling efficiency and effectiveness of coordination in engineering design teams" PhD Thesis Dissertation, Department of Civil Engineering, Stanford University, California, 1993 - Published as DNV Research report 93-2063
Conceptual Interoperability	Wikipedia, Conceptual Interoperability http://en.wikipedia.org/wiki/Conceptual_interoperability
Cuel	A survey on ontology creation methodologies, Cristani, M; Cuel, R. 2005. International Journal on semantic web and information systems.
Data integration	Wikipedia contributors, "Data integration". Wikipedia, The Free Encyclopedia, http://en.wikipedia.org/w/index.php?title=Data_integration&oldid=252572510 (accessed November 28, 2008).
DataGov	Data Governance 2008 Conference. http://www.debtechint.com/dg2008/
Davis	Davis, P.K. and Anderson, R.H. (2003). Improving the Composability of Department of Defense Models and Simulations. RAND Corporation
Davis M	Davis, M. 2006, Semantic Wave 2006 Part 1: Executive Guide to Billion Dollar Markets
DDI	DDI, Data Documentation Initiative. http://www.icpsr.umich.edu/DDI/
DGI	The DGI Data Governance Framework, Gwen Thomas, The Data Governance Institute. http://datagovernance.com/dgi_framework.pdf
Doninini	F. M. Doninini: “Complexity of reasoning”, in /The Description Logic Handbook./ Eds. F. Baader et al. Cambridge University Press 2004.
Dublin Core	Dublin Core, http://dublincore.org/
EIF 2.0	Draft document as basis for EIF 2.0. IDABC. http://ec.europa.eu/idabc/servlets/Doc?id=31597
FFI	Semantic Technologies. FFI Bjørn Jervell Hansen, et al. 22. November 2007. http://rapporter.ffi.no/rapporter/2007/02461.pdf
Galbraith	Jay Galbraith, Designing Complex Organizations – An Information Processing View, 1974, Interfaces, United States
Guarino	An Overview of OntoClean by Nicola Guarino and Christopher A. Welty. http://ontolog.cim3.net/file/resource/presentation/OntoClean--ChrisWelty_20041118/guarinowelty_final_v4.pdf
Gunnlaugsdottir	Using and preserving corporate knowledge during times of change. Johanna Gunnlaugsdottir, Procedings of the 12th Nordic conference for information and documentation, Knowledge and Change. 2004.
Hernández	Mauricio Hernández, Renée J. Miller, Laura M. Haas. Clio: A Semi-Automatic Tool For Schema Mapping. System Demo at SIGMOD, 2001.
Hopwood	Hopwood, Peter (2008-06). "Data Governance: One Size Does Not Fit All". DM Review Magazine. Archived from the original on 2008-10-02.
INTEROP-VLab	INTEROP-VLab: The European Virtual Laboratory for Enterprise interoperability http://interop-vlab.eu/interop-vlab-network/INTEROP-VLab/
ISO 10303	Product model standards http://en.wikipedia.org/wiki/ISO_10303
ISO 1087	ISO 1087, Terminology work, Vocabulary. Part 1: Theory and application

ISO 11179	ISO 11179, Information technology Metadata Registries
ISO 15926	ISO 15926 standard http://www.posccaesar.org/wiki/ISO15926
ISO 15926-2	ISO 15926-2 Data Mode http://www.tc184-sc4.org/wg3ndocs/wg3n1328/lifecycle_integration_schema.html
ISO 15926-RDS	ISO 15926-4Reference Data Library http://www.posccaesar.org/wiki/Rds
ISO 15926-Templates	ISO 15926-7 ongoing work on templates
ISO 2000	ISO 1087-1:2000, Terminology work – Vocabulary – Part 1: Theory and application.
ISO 8000	ISO 8000 Data Quality developed by ISO TC184/SC4
K. Madsen	Enterprise Architecture implementation and management: A case study on interoperability. Kristian Hjort Madsen. IT-University of Copenhagen. IEEE proceedings of the 39th Hawaii international conference on system sciences.
Madsen 2006	Håndbog i begrepsarbeide, Del 2 metoder og forløp. Sundhedsstyrelsen DK 2006. Bodil Nistrup Madsen, DANTERMcentret
MDM conf	MDM SUMMIT 2009 Conference. http://www.tcdii.com/events/cdimdmsummitseries.html
Møller	XML and Document lifecycle. Focus on digital preservation, Master Thesis by K. Helen Møller. Oslo university college, faculty of journalism, library and information science.
Myrseth 2008	Preservation of semantic value - a study of the state of the art. Longrec Project Report. Per Myrseth, Tore R, Christiansen, Helen Gayorfar. 1. april 2008. http://www.longrec.com/Intranet/ResearchResults/StateOfTheArt/Longrec_Understand-SOTA_v1.01.doc
Nardi	D. Nardi and R. J. Brachman: "An intoroduction to description logics", in /The Description Logic Handbook./ Eds. F. Baader et al. Cambridge University Press 2004.
Neuchâtel	Neuchâtel Terminology Model, Classification database object types and their attributes http://www.unece.org/stats/cmf/Part%20I%20Neuchatel_version%202_1.pdf
NHS	National Health Service, Information Governance, http://www.connectingforhealth.nhs.uk/systemsandservices/infogov
Noark	Noark, a specification of functional requirements for electronic recordkeeping systems used in public administration in Norway. http://www.arkivverket.no/english/electronic.html
Obel 2008	Public/Private Partnership. Structuring Toolkit. Peter Obel. Stanford.August 2008.
Ogden	The Meaning of Meaning, C. K. Ogden & I. A. Richards, sixth edition. 1943.
OMG	Object Management Group (OMG™ is an international, open membership, not-for-profit computer industry consortium.) Http://www.omg.org/
OMG BMM	OMGs Business Motivation Model http://www.omg.org/technology/documents/bms_spec_catalog.htm
OMG BPMN	OMGs Business Process Modelling Notation http://www.bpmn.org/

OMG CWM	OMGs Common Warehouse Model http://www.omg.org/CWM
OMG MOF	OMG MOF http://www.omg.org/mof/
OWL	OWL 1.1, http://www.w3.org/Submission/owl11-overview/
Page	Page, E.H., Briggs, R., and Tufarolo, J.A. (2004). Toward a Family of Maturity Models for the Simulation Interconnection Problem. Proceedings of the Spring 2004 Simulation Interoperability Workshop, IEEE CS Press
Perez	Ontological engineering, Asuncion Gomez-Perez et al, Springer 2003.
Price	A Semiotic Information Quality Framework, Rosanne J. Price, Graeme Shanks. The IFIP TC8/WG8.3 International Conference 2004
Provost	On the Cusp: A Global Review of the Semantic Web Industry. David Provost. September 30. 2008. www.davidprovost.com/Resources/Semantic%20Web%20Industry%20Revie.pdf
RDF	RDF http://www.w3.org/RDF/
SAWSDL	SAWSDL Semantic Annotations for WSDL. W3C recommendation 28. august 2007.
Scholl 2007	E-Government Integration and Interoperability: Framing the Research Agenda. Intl Journal of Public Administration, 30: 889–920, 2007. Hans J. (Jochen) Scholl and Ralf Klischewski
Scott	W. Richard Scott, Organizations: Rational, Natural, and Open Systems, 4th Ed. (Upper Saddle River, NJ: Prentice Hall, 1998),
SemanticGov	SemanticGov. Providing Integrated Public Services to Citizens at the National and Pan-European level with the use of Emerging Semantic Web Technologies. http://www.semantic-gov.org/index.php
SemTechTrends	Semantic Web Technologies. Trends and research in ontology-based systems. Editors: John Davis, Rudi Studer, Paul Warren. Wiley 2006.
Sheth	Traveling the Semantic Web through Space, Time , and Theme. By Amit Sheth and Matthew Perry. IEEE internet computing mach/april 2008. http://knoesis.wright.edu/library/download/Sheth-Perry-IEEE08.pdf
Smith-Meyer 2005	Using Models and Model-based Transformation to Generate Operational Context-sensitive Workplaces. Proceedings Collaborative Engineering 2005 H.Smith-Meyer, F.Lillehagen, H.D.Jørgensen, D.Karlsen, E.Dehli, S.Murberg, H.G.Solheim – Troux
Spri 1999	Spri-rapport 481 Metoder och principer i terminologiarbetet. Hälso- och sjukvårdens utvecklingsinstitut, 1999.
Thompson	Thompson, James D. [1967] (2003). Organizations in Action: Social Science Bases of Administrative Theory,
Tolk	Tolk, A. and Muguira, J.A. (2003). The Levels of Conceptual Interoperability Model (LCIM). Proceedings IEEE Fall Simulation Interoperability Workshop, IEEE CS Press
Tolk 2003	Tolk, A. (2003). Beyond Technical Interoperability - Introducing a Reference Model for Measures of Merit for Coalition Interoperability. Proceedings of CCRTS, CCRP Press
Topic Maps	Topic Maps ISO/IEC 13250:2000 "Topic Maps" standard
Turnitsa	Turnitsa, C.D. (2005). Extending the Levels of Conceptual Interoperability Model. Proceedings IEEE Summer Computer Simulation Conference, IEEE CS Press

Veltman	Towards a Semantic Web for Culture, 2004, Veltman. http://jodi.tamu.edu/Articles/v04/i04/Veltman/veltman.pdf
Vlab	Second Scientific Agena of Interop-Vlab, December 10th, 2007. http://interop-vlab.eu/the-scientific-activities/download-the-scientific-agenda-of-interop-vlab/
Wache 2001	Wache, H.;Vgele, T.;Visser,U.; Stuckenschmidt, H.; Schuster, G.; Neumann, H.; Hbner S.: Ontology-Based Information Integration: A Survey. The BUSTER Project, Intelligent Systems Group. Center for Computing Technologies. University of Bremen. 200l.
Yang	Journey to Data quality, Yang W. Lee et al. MIT Press 2006.

TECHNICAL REPORT

- o0o -