# Semantic Interoperability

## for Better Health and Safer Healthcare


**RESEARCH AND DEPLOYMENT ROADMAP FOR EUROPE**

SemanticHEALTH Report
January 2009

European Commission
Information Society and Media

## About the Report

This report was prepared by the SemanticHEALTH project, a Specific Support Action funded by the European Union 6th R&D Framework Programme (FP6). SemanticHEALTH developed a longer-term research and deployment roadmap for semantic interoperability. Its vision is to identify key steps towards realising semantic interoperability across the whole health value system, thereby focusing on the needs of patient care, biomedical and clinical research as well as of public health through the re-use of primary health data.
Partners in the SemanticHEALTH project were: empirica Communication and Technology Research, Germany (coordinator), University of Manchester, United Kingdom, University College London, United Kingdom, WHO Classifications and Terminology, Switzerland, University of Saint Etienne, France, ESKI - National Institute for Strategic Health Research, Hungary, WHO Family of International Classifications Collaborating Centre in the Nordic Countries, Sweden, Radboud University Nijmegen Medical Center, Nijmegen, Netherlands.


## Acknowledgements

The SemanticHEALTH project was partially funded by the European Commission, Directorate General Information Society and Media, through its 6th Framework Programme - a support which is gratefully acknowledged. We thank the colleagues from the 'ICT for Health' Unit for their kind encouragement and support. In particular, we are grateful to Gérard Comyn, Head of Unit, for the support and promotion of this activity, and Ilias Iakovidis, Deputy Head of Unit, for his valuable contributions and continuous encouragement!
We are grateful to numerous global experts and speakers at workshops and sessions organised in the course of the project for sharing their knowledge, experience and critical reviews.
This report reflects solely the views of its contributors. The European Community is not liable for any use that may be made of the information contained therein.





Ref.: Plan-Publi 2009.2098
**Semantic Interoperability for Better Health and Safer Healthcare**
Deployment and Research Roadmap for Europe
Language : EN
Catalogue number : KK-80-09-453-EN-C
ISBN-13 : 978-92-79-11139-6
DOI : 10.2759/38514

A great deal of additional information on the European Union is available on the Internet.
It can be accessed through the Europa server (http://www.europa.eu).





Cover Design & Printed by OIB

# Semantic Interoperability for Better Health and Safer Healthcare

Research and Deployment Roadmap for Europe

Veli N. Stroetmann (Ed.), Dipak Kalra, Pierre Lewalle, Alan Rector, Jean M. Rodrigues, Karl A. Stroetmann, Gyorgy Surjan, Bedirhan Ustun, Martti Virtanen, Pieter E. Zanstra

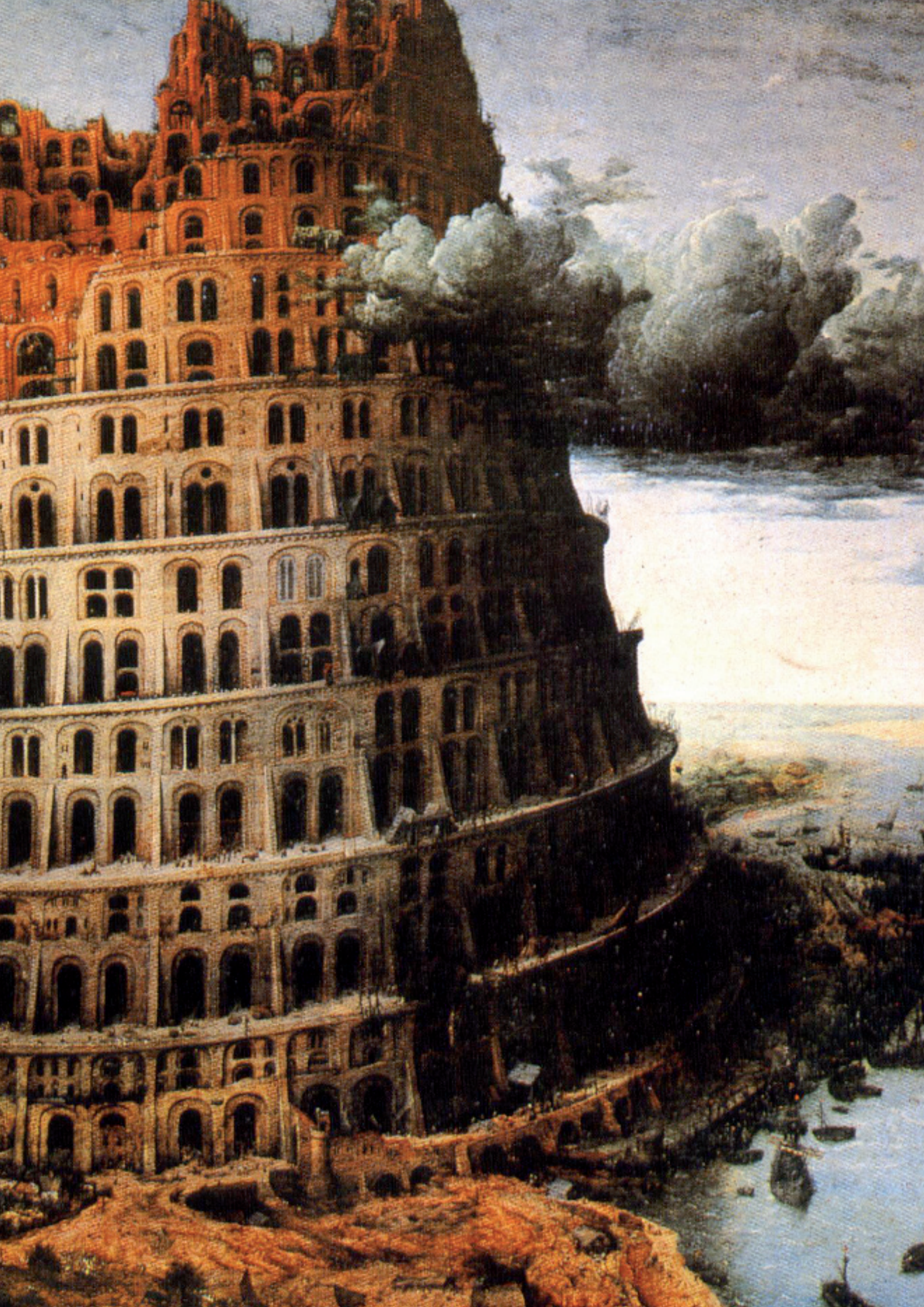
SemanticHEALTH Report

January 2009

Semantic HEALTH

eHealth

European Commission
Information Society and Media



# Foreword

Development and beneficial deployment of electronic health record (EHR) systems has been at the heart of the activities of the ICT for Health (or eHealth) unit of the European Commission since its start 20 years ago. Many R&D projects initiated in the early 90's have by now led to various commercial products and regional networks benefiting European citizens. However, to fully realise the potential of EHR systems we need to ensure a timely and secure access to such systems to all those that are entitled to use them. Moreover, the information contained in EHRs should be up-to-date, accurate and, in its communication to another location, system or language it should be correctly understood. This is called interoperability. Interoperable EHR systems are the most important enabling tools on the road to patient centred care, a lifeline for continuity of care and support to mobility of patients.

The results of the SemanticHEALTH project summarised in this publication underline that interoperability is not just a technical and standardisation challenge. It involves as well leadership, decisions about and investment in political, institutional, organisational, legal and market issues. The European Commission (EC) has long recognised multiple levels of interoperability and addressed the challenges through research, studies, policy documents, awareness raising and support for Member States (MSs) and stakeholders.

In the past several research projects, such as GEHR, SYNAPSES, SYNEX, GALEN and many others, focused on technical issues related to standardised and interoperable EHRs. In recent years, a number of EC-supported studies have focused on interoperability in a wider context: i2-Health developed a conceptual foundation, RIDE drafted a first roadmap, EHR IMPACT analyses benefits of interoperable EHR systems, and three European standards development organisations - CEN, CENELEC, ETSI - have been mandated to develop a work programme for accelerated standardisation in health informatics. Our strategic approach is to focus on the core applications of patient summary and electronic prescribing as "gate-openers" to progress on interoperability across Member States. To further advance this, epSOS (European Patient Smart Open Services), a €22m initiative jointly funded by the EC and 12 national authorities, healthcare providers, industry and other stakeholders, will define, test and validate such applications over a three year period.

The most challenging part remains achieving semantic interoperability of EHR systems. It plays a prominent role in our recently published Recommendation on Interoperability of Electronic Health Record Systems (COM(2008)3282). It calls not only for interoperability at regional and national level but also at EU level - a goal which realistically may take another 20 years to be fully achieved.

Interoperability is about continuous change management. It is a long-term endeavour requiring both permanent structures and the organisation of processes for consensus-building and co-operation among all actors involved. This SemanticHEALTH roadmap helps us to structure the necessary work for many years to come. It underlines that issues of technical standardisation are no longer the most prominent ones in realising the interoperability vision. The benefits from exchanging consistent patient information must become more transparent, confidence must be nurtured that the data will be secure and confidential, organisations must trust those with whom they share information, and the sharing of information can not be seen as in conflict with business or legal interests of the participants. We are sure that the work presented here will contribute to realising this future for the benefit of all people. As that work related to semantic interoperability can be achieved only by full participation and commitment of health professionals, their stronger involvement in all relevant initiatives is a high priority for the EC.

Ilias Iakovidis
Deputy Head of Unit ICT for Health
European Commission, Directorate General
Information Society and Media

# Table of Contents

# 1 Introduction

Real-time eHealth systems integrating all relevant information on a patient as well as medical and other health-related knowledge can not only substantially improve collaborative care, patient safety, quality and efficiency of health services, but also support medical and clinical research, training and public health. The 'holy grail' of healthcare connectivity is a cornerstone for reaping the full benefits of eHealth. However, to fully realise this goal requires interoperability of such systems within health services organisations and jurisdictions, and across regions and countries.

The European Commission has long recognised the need for addressing the multiple levels and complex challenges of interoperability of eHealth solutions. The roots of policy efforts to improve interoperability are grounded in the *European eHealth Action Plan* of 2004[1] and are followed by a number of joint activities with Member States and relevant stakeholders, supported by European projects.[2] The recently published European Commission Recommendation on cross-border interoperability of electronic health record systems aims to contribute to the development of overall European eHealth Interoperability by the end of 2015.[3] The Commission identifies four major levels on which the Member States are encouraged to undertake action. These are the political, organisational, technical and semantic levels, with educational and awareness raising mechanisms to underpin initiatives in those main domains. Semantic interoperability plays a prominent role – it is described as an essential factor in achieving the benefits from electronic health record systems to improve the quality and safety of patient care, public health, clinical research, and health service management.

Building on an extensive review of the latest R&D efforts and in close cooperation with leading experts in Europe and around the world, the SemanticHEALTH project developed a roadmap for further research and deployment of workable solutions in the short and medium term. Aligned in a comprehensive perspective, next steps for realising pragmatic solutions are identified. Three application fields and one cross-cutting domain where action is required on the path to achieving semantic interoperability have been analysed: (1) electronic health records; (2) ontologies and terminologies; (3) public health; (4) socio-economic issues.

A policy of incremental steps at these levels and a focused, modest approach to terminology development is recommended, as well as to undertake a thorough investigation of incentives for their development, implementation, maintenance, and utilisation for implementing the new, patient centred and collaborative health service model. Efforts should focus on optimal - and not 'best' - solutions that are both realistic and affordable. Four priority areas and related challenges that would benefit most from these developments have been identified:

- *Patient care:* patient safety; dissemination of good practice, integration of education and care; connecting multiple locations for collaborative care delivery (at local, regional, na-tional and international levels); empowerment of citizens (patient centred healthcare)
- *Public health:* international statistics; comparative outcome assessment; pharmacovigilance; coordination of risk assessment, management and surveillance of large-scale adverse health events, population health research
- *Research and translational medicine:* multi-centre studies and trials, health data repositories, bio- and tissue-banks, development of personalised medicine based on genetic and genomic analyses
- *Support for diverse markets:* identification of solutions with superior benefit/cost ratios; enabling plug-and-play best of breed, encouraging industry involvement, especially SMEs; stimulating innovations by health service providers and involving clinicians, harmonising legal and regulatory frameworks.

In this concise report, major conceptual innovations of the project and key recommendations for the further development of semantic interoperability are outlined.

Firstly, a brief summary of the definition of semantic interoperability and the analytical framework of the project is provided. Secondly, the vision for semantic interoperability which guides all further recommendation work is presented. Thirdly, some key business use cases for semantic interoperability are summarised. Then a contextualisation of the whole phenomenon in its socio-political, legal and economic context follows. The final chapter presents the recommendations that follow from the preceding chapters and concludes with an outlook on future developments.

Based on the foundations laid by recent research, Europe has a unique window of opportunity to make faster progress in this area because many of its health system stakeholders are very well aware of the cultural and linguistic challenges the new model of collaborative healthcare involves, but also of the opportunities arising to meet the growing needs for the sustainability of our health systems. Most Member States are presently investing in eHealth strategies and infrastructures, and the European Commission can be expected to provide continuing help and coordinating mechanisms to support them.

The scientific foundations of this work and many more details can be found in the public deliverables of the study available at *www.semantichealth.org*.

1. COM(2004) 356(final): e-Health – making healthcare better for European citizens: An action plan for a European e-health area : http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:52004DC0356:EN:NOT
2. Related EC-supported projects dealing with interoperability issues are: RIDE (www.srdc.metu.edu.tr/webpage/projects/ride), i2-Health (www.i2-health.org), EHR-Impact (www.ehr-impact.eu), Q-Rec (www.eurorec.org/projects/qrec.cfm), epSOS (www.epsos.eu), CALLIOPE (www.calliope-network.eu), etc.
3. COM(2008)3282 final: Commission Recommendation on cross-border interoperability of electronic health record systems, Brus-sels, 2008-07-02, http://ec.europa.eu/information_society/activities/health/docs/policy/20080702-interop_recom.pdf

# 2 Defining semantic interoperability

The SemanticHEALTH study applies the following overall interoperability (IOp) definition: *Health system interoperability is the ability, facilitated by ICT applications and systems,*

- *to exchange, understand and act on citizens/patients and other health-related information and knowledge*
- *among linguistically and culturally disparate health professionals, patients and other actors and organisations*
- *within and across health system jurisdictions in a collaborative manner.*

In this context, semantic interoperability (SIOp) addresses issues of how to best facilitate the coding, transmission and use of meaning across seamless health services, between providers, patients, citizens and authorities, research and training. Its geographic scope ranges from local interoperability (within, e.g., hospitals or hospital networks) to regional, national and crossborder interoperability. The information transferred may be at the level of individual patients, but also aggregated information for quality assurance, policy, remuneration, or research.[4]

Semantic interoperability (SIOp) has numerous facets:

- *For individual patients* SIOp relevant tasks comprise assisted clinical data capture and quick access to the patient record as well as to pertinent background knowledge. It also includes quality assurance, clinical decision support, monitoring and alerts, as well as feedback regarding quality and costs.
- *For aggregated population data* SIOp relevant tasks include reporting, health economics, surveillance, quality assurance, epidemiology (hypothesis formulation), bio- and tissue-banking.
- *SIOp* enables the meaningful linkage of research findings and knowledge to patient information, and the discovery of new knowledge from semantically coherent EHR repositories.
- In addition to precision of meaning, consistancy, understandability and reproducibility are three major desiderata for semantically interoperable systems:
  - *Consistancy* means that the receiving system must be able to recognise what has been sent, so it is a prime requirement for machine-machine communications and dictates the need for unambiguous identifiers.
  - *Understandability* is essential for human communication. Humans can tolerate considerable ambiguity, but tend to focus too narrowly, so that the requirements are almost the reverse as for automated support. It is limited by the trust that the information is valid, especially with aggregated population data where the aggregation process may result in loss of information.
  - *Reproducibility* addresses the question of inter-individual reliability when data are collected or encoded. This holds both for individual and aggregated data.

It is desirable that semantic interoperability will be achieved gradually. E.g., where considered useful (e.g. for patient safety), at modest cost clinical terminologies will increasingly pervade public health, good practice in healthcare and other applications.

In essence the *Semantic*HEALTH goal is to work towards and support collaboration among human actors and stakeholders, rather than only interoperability among computers.

To further clarify these issues, the research distinguishes four levels of IOp, two of them relating to semantic interoperability:

| | |
|---|---|
| Level 0: | no interoperability at all |
| Level 1: | technical and syntactical interoperability (no semantic interoperability) |
| Level 2: | two orthogonal levels of partial semantic interoperability<br>*Level 2a:* unidirectional semantic interoperability<br>*Level 2b:* bidirectional semantic interoperability of meaningful fragments |
| Level 3: | full semantic interoperability, sharable context, seamless co-operability |

To explain and distinguish the 4 different levels, consider the following scenario: 56 year old Pádraig recently moved from Ireland to Spain to take up his new job in a multinational IT company. A few weeks after arriving, he falls ill, consults his local (Spanish) GP and is transferred to the next hospital for further tests. Depending on the level of SIOp established, the hospital has to initiate the following steps:

- Level 0 (no interoperability at all): Pádraig has to undergo a full set of lengthy investigations for the doctors to find out the cause of his severe pain. Unfortunately, results from the local GP as well as from his Irish GP are not available at the point of care within the hospital due to the missing technical equipment.

- Level 1 (technical and syntactical interoperability): Pádraigs doctor in the hospital is able to receive electronic documents that were released from the Irish GP as well as his local GP upon request. Widely available applications supporting syntactical interoperability (such as web browsers and email clients), allow the download of patient data and provide immediate access. Unfortunately, none of the available doctors in the hospital is able to translate the Irish document, and only human intervention allows interpreting the information submitted by the local GP for adding into the hospitals information system.

- Level 2 (partial semantic interoperability): The Spanish hospital doctor is able to securely access via the Internet parts of Pádraig's Electronic Health Record released by his Irish GP as well as the local GP that he visited just hours earlier. Although both documents contain mostly free text, fragments of high importance (such as demographics, allergies, diagnoses, and parts of medical history) are encoded using international coding schemes, which the hospital information system can automatically detect, interpret and meaningfully present to the attending physician.

- Level 3 (full semantic interoperability, co-operability): In this ideal situation and after thorough authentication took place, the Spanish hospital information system is able to automatically access, interpret and present all necessary medical information about Pádraig to the physician at the point of care. Neither language nor technological differences prevent the system to seamlessly integrate the received information into the local record and provide a complete picture of Pádraigs health as if it would have been collected locally. Further, the anonymised data feeds directly into the tools of public health authorities and researchers.

The partial nature of SIOp may be expressed in terms of a part-total ratio. For example, there may be SIOp within a number of institutions, but lack of SIOp across them.
In other words, SIOp might not exist as an all-pervasive state, but rather be a description of the relationship between specified systems or services.

It must be kept in mind that SIOp implementation also depends on social, cultural and human factors within each organisation, region and country, each system and each time period. Realising full SIOp is not necessarily a consensual goal in every place at any fixed time.

4. The EC Recommendation, COM(2008)3282 final, applies the following definition: "Semantic interoperability means ensuring that the precise meaning of exchanged information is understandable by any other system or application not initially developed for this purpose", whereas "interoperability of electronic health record systems means the ability of two or more electronic health record systems to exchange both computer interpretable data and human interpretable information and knowledge", p.14.
http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:32008H0594:EN:NOT

# 3 Semantic interoperability, the goal, the vision and the challenges

## 3.1. The SIOp goal

Health professionals of all disciplines require access to detailed and complete health data and patient records in order to manage the safe and effective delivery of health services. These records need to be linked to salient domain knowledge and guidance, and to be shared in real time within and between care teams across geographical and linguistic boundaries. Patients and their families often also require access to their own health records, to suitable educational materials, and also to specialized medical content such as scientific literature and clinical practice guidelines. This permits them to play an active role in their health management, in partnership with health professionals.

The clinical interoperability requirement from health IT is for clinical meaning to be expressed consistently within electronic health record (EHR) systems and medical knowledge repositories, in particular where computers in addition to humans need to be able to process such data safely. This is particularly needed if computational services are to be able to interpret safely clinical data that has been integrated from diverse sources.

In the specific case of EHR systems and their semantic interoperability requirements, we need:

- to enable the safe, meaningful sharing and combining of health record data between heterogeneous systems;
- to enable the consistent use of modern terminology systems and medical knowledge resources;
- to enable the integration and safe use of computerised protocols, alerts and care pathways by EHR systems;
- to link EHR data to explanatory and educational materials to support patient and family engagement and professional development;
- to ensure the necessary data quality and consistency to enable rigorous secondary uses of longitudinal and heterogeneous data: public health, research, health service management.

Interoperability requires agreement on meanings and labels for those meanings – on ontologies and lexicons, which together we label as terminologies. The primary goal of ontologies and terminologies for interoperability is to enable the faithful exchange of meaning between machines and between machines and people.

## 3.2 Vision for the future

The 'hopeful' vision presented here is the result of long discussions with experts, meetings with industry representatives and feedback gathered through workshops and seminars. It is important to note that, regardless of the type of vision one may develop, semantic interoperability is not a phenomenon to be expected over night. Rather, it will emerge gradually and may even in the most optimistic of assumptions remain an incomplete phenomenon. The SemanticHEALTH vision is characterised by a large number of changes at both the technical and the use case level. Note however, that even in this vision, no full semantic interoperability or a complete harmonisation of either EHR models or terminologies can be expected.

### 3.2.1 TECHNICAL EVOLUTION

Semantic operability will be achieved only gradually beginning with applications with high benefit and modest cost. Given appropriate incentives, there will be a series of bottom up and top down measures that will achieve a level of interoperability that protects patient safety and

supports common undertakings in public health, clinical research, and dissemination of good practice. Material for dissemination of good clinical practice will increasingly be linked to the structures and terminologies used for clinical care. Nevertheless, semantic interoperability will not be complete. Much healthcare will continue to be delivered locally using idiosyncratic systems or with minimal, or no, IT support. Nor will there be complete harmonisation of either EHR models or terminologies. There will continue to be a major requirement for mappings and transformation services based on technologies analogous (or identical) to current data warehousing and mediation technologies.

For terminologies, this will best be achieved by starting with areas where there is a high degree of consensus on both the content and the need. Key areas are likely to be sensitivities and adverse drug reactions, translational medicine, and large scale public health and population research initiatives such as "biobanking".

The development mechanisms that are successful will be open, collaborative and Web-enabled, and specialised communities will contribute significantly to the effort by standardising vocabularies for local purposes. These communities will 'own' and take responsibility for their terminologies, helped by central servers and technologies which they will think of as part of their environment, just as much of the population today thinks of the Web, Google, Facebook, Flickr and related applications as just 'there'.

The methods will become increasingly formal. The conflict between the scaling problems presented by terminologies with pre-coordinated terms and the difficulty of maintaining consistency with compositional terminologies (that assemble terms as they are used) will be overcome. To this end, the formal structure of terminologies must be well-devised to take advantage of logic-based underpinnings. The same applies to clinical information models that need to take advantage of modern technologies to ensure their mutual consistency and consistent binding to the new terminologies. Common links to terminologies used in molecular biology will be forged.

Because of obvious usefulness, there will be serious involvement by clinical staff in medical terminologies as there is by bioinformatics and molecular biologists in bioontologies.

### 3.2.2 EVOLUTION WITH RESPECT TO APPLICATION FIELDS

Patient care will benefit dramatically with a significant reduction in avoidable errors and improvements in patient safety. Distributed care will become the dominant paradigm, with a rapid shift of care both to the community and to highly specialised centres applying the latest techniques arising from accelerated clinical and translational research. Care in remote areas will be particularly affected. Patients will take increasing responsibility for their own care with the help of Web-enabled tools that link directly both to their own records and to the records held in the various institutions in which they seek care. The elapsed time to translate new findings into practice will be drastically reduced. The rise in the overall cost of care will be mitigated.

Public health will be facilitated by much faster and less costly collection of regional, national and international statistics, as most statistics will be derived from data collected during patient care, although there will remain a need for experts to monitor and check data for critical measures. Surveillance for the emergence of new epidemic diseases and major health problems will become more effective, and most outbreaks will be recognised early enough to be contained, although the increasing population and rate of travel will result in more small outbreaks.

Clinical and translational research will advance very rapidly. Information sharing amongst researchers will be the norm. The lines between patient care and translational research, as well as between translational research and basic research in molecular biology will become increas-

5. http://demo.mapofmedicine.com/

ingly blurred. Most studies will be large scale, international. Many will reuse data from earlier studies to triage hypotheses and minimise the number of patients exposed to unsuccessful therapies. Research will depend increasingly on bio- and tissue banks which will have access to rich information on the lifelong outcome of large cohorts of patients collected in the course of their routine care. A uniform structure of privacy, consent and governance will manage data sharing for research in ways that are accepted by the vast majority of the population.

A balanced market will develop with large suppliers managing hospitals as a whole but with innovative SMEs and specialist vendors supplying systems to address special functions and niche markets. The evolution between large and small, institutional and personal suppliers will be fluid, and European companies will play a major part in the overall commercial market. The time required to integrate a new specialised module or system into a hospital's infrastructure will drop from person years to person weeks, in some cases to person-hours. The difficulty of integrating systems will cease to be a barrier to adoption of best-of-breed solutions, and they will be embraced by central administrations and central IT departments.

## 3.3 Key trends

This section presents 'inevitable' outcomes in the form of a simple extension of trends that are already visible today, most notably in the form of advanced statistical text and web mining technologies. Also with regard to application fields, a more active citizenry and increasingly expensive treatments are 'inevitable' outcomes, given our knowledge about today's trends.

### 3.3.1 TECHNICAL TRENDS

Statistical text and Web mining technologies will advance rapidly, and Google-like technologies will take over much of the burden of coarse grained search for navigation information discovery. This will probably include linking of EHR systems to text material for decision support such as the *Map of Medicine*[5]. A balance between semantic and statistical technologies will eventually be established, but where the balance will be remains to be seen. Cross language searching will improve rapidly, driven by general commercial imperatives, but is unlikely to eliminate the need for multilingual systems. Research on how best to use the two in concert is a major priority.

*Direct encoding of free text into formal* vocabularies and EHR structures will improve radically, partly driven by voice recognition.

*Personal medical systems* will proliferate. Whether they interact effectively with the local health care systems will depend on a combination of technical and commercial pressures. They may become a key driver for interoperability or may operate entirely outside it.

Concerns about *privacy and confidentiality* will continue to be key limiting factors in interoperability, and may impede developments that would be technically feasible and beneficial.

### 3.3.2 TRENDS WITH RESPECT TO APPLICATION FIELDS

Patients will increasingly use web resources and take responsibility for their own care, with or without coordination with professional carers.

Clinical medicine and medical technologies will advance and new treatments will inevitably be more expensive. The breadth and diversity of medical, life science, psycho-social and environmental knowledge and their relevance for health will continue to expand rapidly.

Developing and integrating ontologies into decision support and knowledge management software will become a key priority.

The role of quality control, benchmarking and ranking, economic efficiency as well as societal benefits will increase further.

## 3.4 Challenges

Major challenges to be aware of could be the result of a "do-nothing" or "do-too-much" development path. On the technical level, a lack of incentives or, alternatively, a sudden overspending on semantic interoperability may both be harmful. With regard to application fields, a continuation on the status-quo path will lead to impediment of research and to stifled competition in the healthcare industry. The following observations therefore are intended to strongly underline the urgent need for appropriate action.

### 3.4.1 TECHNICAL CHALLENGES

Without greater economic or policy/regulatory *incentives* to interoperability little will be done and the status quo may be maintained. Semantic interoperability will remain confined to special cases, with little advance on the current state. National and specialist terminologies and EHR formats will remain silos. Virtually all records for patient care will remain in free text.

Alternatively, enormous resources may be spent on *over-ambitious* plans for semantic interoperability which will inevitably fail. In either case communication will take place by going around rather than via the clinical information systems. In countries where it is mandated, large and unwieldy approaches such as SNOMED CT and HL7 V3 will become taxes on healthcare, absorbing significant

resources while returning no, or in some cases even negative, benefits.
*Terminologies* will remain closed or partly closed. Most of the developing world and much of Europe will be excluded from their development, so they will neither fit their purposes nor be owned by them. The revision time for major terminologies will remain years. The release time for "mandated" terminologies will remain months.

Without major changes, *obsolete technologies and primitive tools* will prevail without formal validation and support. The defects in the resulting systems will be large enough that no one will trust them. Resources will be spent for workarounds, so that the legacy becomes increasingly difficult to change.

The profession will remain alienated from informatics in general and `coding' in particular.

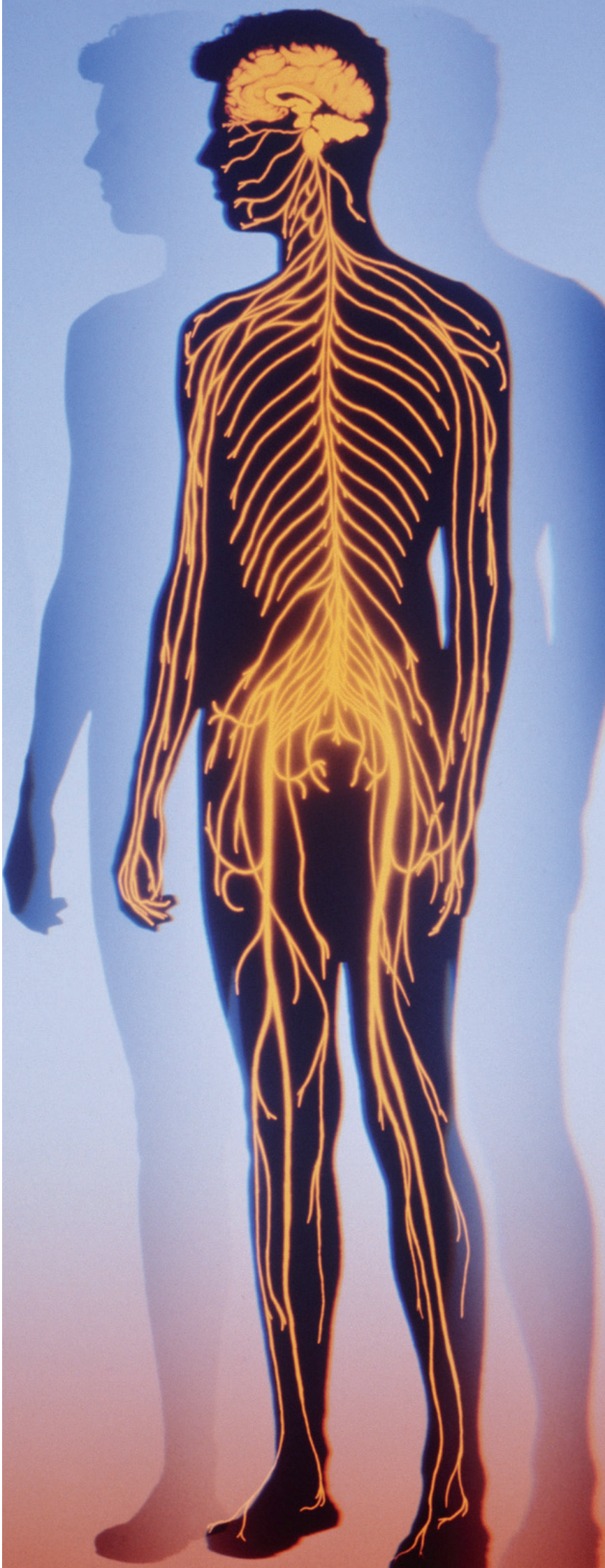
### 3.4.2 CHALLENGES OF APPLICATION FIELDS

Without active policy interventions and further research, there will be tardy progress in patient safety improvement and reduction in clinical errors will (continue to) be slow and sporadic. The current global death toll of hundreds of thousands per annum will continue. Good practice will (continue to) require up to twenty years to be adopted by the profession as a whole, even when evidence in its favour is unequivocal. Healthcare will continue to be fragmented institutionally. Errors resulting from miscommunication will (continue to) occur and account for significant morbidity. Well educated and informed patients may take matters into their own hands, but those with fewer resources or in marginalized areas will be left at the mercy of chance and hazard.

*Public health* will (continue to) depend on specialist encoders and be limited by the cost and accuracy of capturing information post-hoc. Bio-surveillance will remain a specialist activity divorced from mainstream clinical practice.

*Clinical and translational research* will continue to be conducted in silos. The cost of mounting multicentre trials will become the dominant barrier in the application of basic biological knowledge to medical care.

*The market* will (continue to) be dominated by a few large suppliers who supply `complete', one-size-fits-all solutions to entire hospitals or even countries. Innovation will become more difficult. Niche systems will be rare and will not interact with the main systems. Integration related difficulties will be the major barrier to the procurement of specialist systems which will be resisted vigorously by administrative and central IT directorates.



# 4 SemanticHEALTH roadmap and recommendations

Building on an extensive review of the latest R&D efforts in Europe and around the world[6] and various workshops and exchanges with global experts in the field, a roadmap as well as detailed key recommendations for further research and deployment of workable solutions in the short and medium term are now presented. To help structure the overall domain of semantic interoperability and focus on areas of particular relevance in the context of European Union eHealth efforts and research programmes, three application fields and one cross-cutting domain where identified where priority actions are required on the path to achieving semantic interoperability: (1) Electronic health record systems; (2) Ontologies and terminologies; (3) Public health; (4) Socio-economic issues.

## 4.1 Electronic health record systems[7]

### 4.1.1 THE NEED FOR AND BENEFITS FROM SIOP

Semantic interoperability is most needed when electronic health record (EHR) data are to be shared and combined from different systems (or across diverse modules within a large system). Full semantic interoperability (Level 3) is required across heterogeneous EHR systems in order to gain the benefits of computerised support for reminders, alerts, decision support, workflow management and evidence-based healthcare, i.e. to improve effectiveness and reduce clinical risks. The key semantic interoperability requirement identified in support of evidence based and safe clinical care is the ability to search for particular EHR data entries that are of relevance to such functionalities.

It is recommended that such a high level of SIOp is initially only sought in specific areas of clinical practice that are known to be of high patient safety relevance, and in priority areas for which the evidence is strongest for a gap to be bridged between current and good practice. The way forward in tackling the challenges of semantic interoperability is to prioritise and focus on tractable subsections of the EHR knowledge domain. Current attempts to standardise the capture, representation and communication of clinical (EHR) data reply upon three layers of artefact to represent meaning:

1. Generic reference models for representing clinical (EHR) data, e.g. ISO/EN 13606 Part 1[8], HL7 CDA Release 2[9], the openEHR Reference Model[10]
2. Agreed clinical data structure definitions, e.g. openEHR archetypes[11], ISO/EN 13606 Part 2[12], HL7 templates[13], generic templates and data sets
3. Clinical terminology systems, e.g., LOINC[14] and SNOMED CT[15]

In selecting a reference model for EHR interoperability, the main global candidates are ISO/EN 13606 Part 1, openEHR (both of which are optimised to work with clinical archetypes) and HL7 Clinical Document Architecture (which is limited to a single document). Although IHE's Cross Document Sharing approach (XDS) is a useful stepping stone towards a full EHR solution, its semantic support is presently much weaker, being limited to a small number of document indices (metadata) and, although useful, is not capable of supporting Level 2 or Level 3 semantic interoperability. In selecting the clinical data structure definitions archetypes and templates are the main candidates. An archetype instance is a knowledge artefact that defines how the EHR reference model hierarchy should be organised to represent the data for one clinical entry or care scenario. Because these archetype definitions are represented in a standardised form, they can be shared and used across record-sharing communities to define how locally-organised clinical data should be mapped consistently (even if the data originate from multiple systems). HL7 Templates serve a slightly different purpose, as a means to constrain and verify conformance to profiled HL7 Version 3 Refined Message Information Models (RMIMs). A template is an expression of a set of constraints on the RIM which is used to apply additional constraints to a portion of an in-

stance of data which is expressed in terms of some other Static Model. Templates are used to further define and refine these existing models within a narrower and more focused scope.[16]

On the one hand, full semantic interoperability cannot be reached without a clear sharing of roles between reference model, archetypes structure and terminology which are all necessary. On the other hand, when one of the components is claiming its ability to produce full semantic interoperability alone or under the condition that the two other components conform to its needs, as it has been proposed very often in the past and still in the present, then the goal of full semantic interoperability cannot be reached.

Sharing clinical meaning does not automatically imply (and cannot require) identical terms and data structures: different physical or logical EHR representations may have a common meaning i.e. they may be semantically equivalent. Therefore the goal of semantic interoperability is: to be able to recognise and process semantically equivalent information homogeneously, even if instances are heterogeneously represented, i.e. if they are differently structured, and/or using different terminology systems, and/or using different natural languages. This equivalence needs to be robustly computable, and not just human readable, in order for guidelines, care pathways, alerting and decision support components to function effectively and safely across EHRs that have been combined from heterogeneous systems.

From the perspective of the EHR, achieving Level 1 (syntactic interoperability) enables the exchange of health record information to an extent that permits the mapping of corresponding parts of an information structure between systems, so that data for the relevant patient can be imported and can be selected and retrieved according to non-semantic properties such as the date of recording or the originating provider, and also searched by some coarse grained semantic categories such as a document type. This kind of interoperability is achieved, for example, by using a standard EHR reference model (without any semantic structures such as archetypes) or by using the IHE[17] XDS (cross-document sharing) profile. It can support clinical shared care, in which the main requirement is for the human readability of documents organised by date and for a modest degree of filtering via coarse grained properties.

6. See SemanticHEALTH D7.2: Semantic Interoperability Deployment and Research Roadmap), www.semantichealth.org
7. Dipak Kalra: Barriers, approaches and research priorities for semantic interoperability in support of clinical care delivery – EHR, SemanticHEALTH D4.1, EC-supported FP6 SA, December 2007
8. Kalra D, Lloyd D. EN13606 Electronic Health Record Communication Part 1: Reference Model. CEN TC/251, Brussels. Febru-ary 2007
9. Dolin R et al. HL7 Clinical Document Architecture Release 2.0. Health Level 7, May 2005
10. Beale T, Lloyd D (editors). The openEHR Reference Model version 1.0.1. Available from http://svn.openehr.org/specification/ TAGS/Release-1.0.1/ publishing/index.html (last accessed April 2007)
11. Beale T (editor). The openEHR Archetype Model (AOM) version 1.0.1. Available from http://svn.openehr.org/specification/ BRANCHES/Release-1.0.1-candidate/publishing/architecture/am/aom.pdf (last accessed April 2007)
12. Kalra, Beale T, Heard S, Lloyd D. EN13606 Electronic Health Record Communication Part 2: Archetype Interchange Specifica-tion; CEN TC/251, Brussels. 2007
13. Grieve G, Hamm R, Shafarman M, Mulrooney G. HL7 Template Specification. Health Level 7, 2007
14. Logical Observation Identifiers Names and Codes (LOINC). Please see http://www.regenstrief.org/loinc/
15. SNOMED CT - Systematized Nomenclature of Medicine Clinical Terms. Please see http://www.ihtsdo.org/our-standards
16. A specific collaboration between the EN 13606 archetype development team and HL7 Templates has resulted in a significant commonality of metadata associated with these artefacts, and some alignment of the constraint functions supported by each. However, the use cases were considered to be too different for the archetype formalism to be adopted directly by HL7.
17. See http://www.ihe-europe.net/ (integrating the healthcare enterprise initiative)

Level 2 (partial semantic interoperability) can be achieved in one of two ways. Level 2a (unidirectional semantic interoperability) is achieved by using a deeper level of data structure than simple documents and headings, i.e. finer grained entries are structured and labelled, but in ways determined by each system or vendor. A mapping is required in order for the receiving system to correctly match imported data items with the corresponding equivalents in the local repository. There may be a poor alignment of data values, resulting in the need for a code mapping (translation) process to occur either at import or whenever data are queried and retrieved. This kind of interoperability is often required and provided by interface tools when data are migrated from legacy to new systems. This is often a costly process, as many complex mappings and data transformations have to be defined and implemented. This approach is usually used to integrate whole repositories, but is too costly to implement for the transfer of individual patient records between ad hoc systems for shared care.

Level 2b (semantic interoperability of meaningful fragments) is attained by agreeing and sharing fine grained data structures between sender and receiver, as has historically been carried out via predefined clinical messages (e.g. for screening and immunisation programmes, claims reimbursement) and is now being adopted for the transfer of prescription information within eHealth programmes. In these cases the mappings are performed per data set for the import of a standard message and not to attempt to import data from diverse source-system structures. This approach is scalable only to the extent that a library of predefined standard messages can be developed and maintained. Most national systems can manage 50-100 of these, but not many hundred.

In Level 3 (full semantic interoperability) the use of an EHR reference model, a rich library of clinical data structures, and the definitions of terminology bindings to value lists for each element of the data structures have all to be agreed within a record sharing community. This permits any arbitrary extracts from EHRs to be imported and combined with locally-held data seamlessly without the need for specific mappings. New data structure definitions can periodically be defined and published as required (e.g. as archetypes) and readily leveraged by existing systems with minimal effort. Queries can be distributed and executed on heterogeneous repositories and their result sets combined, so that there is no longer any functional difference between a centralised or a federated EHR.

Levels 1, 2a and 2b are already achieved to varying extents in different countries, and different care settings. They can confer benefits, enabling the support of predefined messages and clinical shared care in situations where the data set is discrete and modest in size, or where human readable records are primarily needed. Much improvement in today's quality and continuity of care across boundaries can be achieved by working towards level 2b interoperability for specific aspects of health care such as the management of long-term conditions (e.g. chronic heart failure, diabetes, asthma) where good clinical consensus already exists on the core data items that should be collected and shared.

### 4.1.2 PRIORITY EHR APPLICATION FIELDS AND RECOMMENDATIONS

However, the investigations undertaken through SemanticHEALTH amongst a wide range of international stakeholders suggests that there is now a growing concern to address the more complex and generalised challenges of patient safety and the cost-effective and equitable use of healthcare resources. The findings of this project suggest that full semantic interoperability (Level 3, as defined by SemanticHEALTH) is required across heterogeneous EHR systems in order to gain the benefits of computerised support for reminders, alerts, decision support, workflow management and evidence-based health care, i.e. to improve effectiveness and reduce clinical risk. However, it is recognised that achieving Level 3 across the entirety of healthcare would be a lengthy, expensive and possibly unattainable goal.

It is instead recommended that Level 3 interoperability is sought in specific areas of clinical practice that are known to be of high patient safety risk, and in priority areas for which the evidence is strongest for a gap to be bridged between current and good practice. In effect, these are the cases for which computerised decision support and care pathway support are most needed.

These priority areas are:

- New medication prescriptions requiring comprehensive information on concurrent medication and details of known allergies and conditions (not simple ETP – Electronic Transfer of Prescription)
- Reminders and prompts for overdue or overlooked health care actions and interventions
- Evidence-based care, the use of clinical guidelines and other forms of evidence to determine the optimal management strategy and care pathway for a given patient
- Care transfers, referrals and within-team workflow prompts such as the degree of urgency and the expectations of the referring clinician from another team member
- Care coordination ensuring that a high-level view can be taken of distributed (multi-team) care to protect against duplication, delay and incompatible interventions.

Many eHealth programmes have begun this challenge by focussing on patient summaries, although often without first defining the target business use cases to be supported by them. The greatest immediate and most urgent benefit to be realised from a medical summary is the improved safety of new medication prescriptions, including the ability to leverage decision support. eHealth should therefore target the summary towards that data content that will inform this kind of decision support: current and recent previous medication, known allergies, and those clinical conditions that are known to be contraindications to some medication items (not necessarily requiring comprehensive multiprofessional problem lists).

The following short and medium term tasks on the road to enable EHR system SIOp in the future have been identified:

*a) Areas needing adoption (short term actions):*
The focus is here on national eHealth programmes and thereby also on industry, and ideally a pan-European perspective should be chosen:

- Agree on a generic model for EHR communications: consider the adoption and promotion of EN13606
- Adopt a standardised approach for representing and sharing of clinical data structure specifications: agree to use archetypes
- Collaborate on key use cases for shared care and patient safety, and on defining and tidying the corresponding SNOMED CT subsets
- Develop and share policies on SNOMED CT term co-ordination
- Seed clinical fora to develop care pathways and archetypes to meet the needs of safe and evidence based care in different medical domains and disciplines
- Strengthen clinical user training in the use of EHRs, terminology and structured records.

*b) Areas needing wide-scale evaluations (medium term):*
Here some results exist, but need refinement and real clinical use, to determine good practice:

- Develop good practice in archetype design and terminology binding to them
- Formalise the governance and quality labelling of archetypes and other knowledge resources
- Establish useful exemplars of SNOMED CT subsets being adopted within EHR systems and delivered in meaningful ways to clinical users
- Develop the business rules and validation processes to support term coordination (pre- and post-coordination).

*c) Areas needing investment (medium term)*
The business cases are not yet strong enough for industry, but products are needed: maybe sponsored open source initiatives:

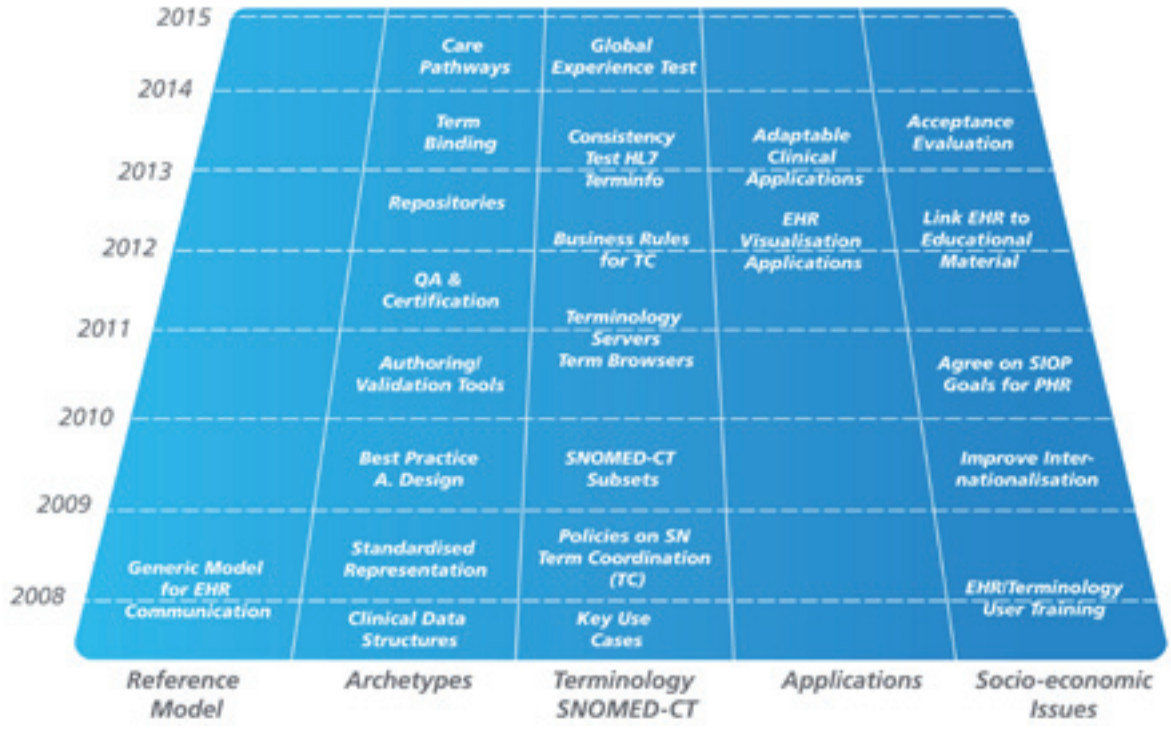
- Archetype & template authoring and validation tools
- Terminology servers and term browsers for SNOMED CT, including support for term coordination.

*d) Areas needing further (focused) research (long term)*
These are also identified for consideration and support in future EU Framework programmes:

- Quality assurance and certification of archetypes
- Archetype indexing, ontology binding to archetypes, and archetype/template repository services
- EHR visualisation applications that can support search and navigation within large and complex health record systems electronically
- Linking EHR data to educational materials and clinical evidence, to enable consumer engagement and support health professional training
- Semantic interoperability goals and solutions for Personal Health Records and near patient eHealth
- Adaptable clinical applications that can reflect evidence-based data structures
- Investigation of knowledge management resources necessary to foster records in which all entries are fully computable
- Test the HL7 TermInfo Trial Standard for consistent implementation and usage, further test SNOMED CT for more global experience.

This roadmap for Electronic Health Records is reproduced and summarised in Figure 1.

FIGURE 1: SUMMARY ROADMAP FOR ELECTRONIC HEALTH RECORDS



To improve the quality of patient care, the following specific areas need to be addressed:

- Patient safety
  - reduce avoidable errors
  - coordinate increasingly complex care [«manage» --> «coordinate»]
  - foster evidence based care
  - monitor good practice
  - reduce duplication and delay
- Disseminate good practice, integrate education and care
- Link patient records to guidelines and care pathways
- Link patient records to professional and patient educational resources
- Connect multiple carer locations of delivery (at local, regional, national and international levels)
- Support team-based care
- Coordinate care between multiple specialist centres and primary care
- Empower and involve citizens - patient centred healthcare

# 4.2 Terminologies and ontologies

### 4.2.1 DEFINITIONS

The vocabulary used to describe terminologies, ontologies, and classification systems has always been a source of confusion, since different authors have used the same words differently. Although the use of the term 'ontology' has proliferated in the research community, there is some reason to cast doubt on the claims made on ontology's behalf: Too many recent publications, calls for research proposals and project descriptions have nurtured what are in our view insupportable expectations. It is thus understandable that some have been tempted to see in ontology just one more new catchword. Another problematic term is "knowledge" which tends to be used in an inflationary way without clarifying the issues it wants to address.

To start with, we therefore provide below a list of common definitions used by us:[18]

- *Controlled Vocabulary* – a list of specified items to be used for some purpose, usually in an information system to reduce ambiguity, misspellings, etc.
- *System of identifiers* ("codes") – Controlled vocabularies, and many lexicons, ontologies, and thesauri, are usually accompanied by systems of identifiers for their units, e.g., identifiers act as the primary unambiguous means of referring to the entities in the system for computational purposes with the text form being used for communication with users. Examples are the "Concept Unique Identifiers (CUIs) from the Unified Medical Language System (UMLS), SNOMED CT IDs.
- *Lexicon* – A list of linguistic units that may be attached to a controlled vocabulary or ontology, in a specific language or sublanguage, often including linguistic information such as synonyms, preferred terms, parts of speech, inflections and other grammatical material. Example: Term terms and lexical material in UMLS identified by Lexical Unique Identifiers (LUIs).
- *Ontology* (sensu information system) – a symbolic logical model of some part of the meanings of the notions used in a field, i.e. those things which are universally true or true by definition.[19] The key relationship in an ontology is "subsumption" or "kind-of". Every instance of a subkind must be an instance of the kind, without exception. Typically ontologies are implemented in logic languages such as Ontylog or OWL or frame systems such as Protégé-Frames. Examples: The GALEN Core Model, the stated form of SNOMED CT.
- *Classification* – an organisation of entities into classes for a specific purpose such as international reporting or remuneration. Examples ICD and Diagnosis Related Groups.
- *Thesaurus* – a system of terms organised for navigation with the primary relationship being "broader than"/ "narrower than". Examples MeSH, WordNet.
- *Background knowledge base* – or "*Knowledge Representation System*" – the common knowledge to be assumed by the system, including both the ontology – what is universally true – and generalisations about what is typically true.
- *Terminology* – Any or all of the above in various combinations. Most heath terminologies consist, at a minimum, of a controlled vocabulary and a system of identifiers. They may include extended lexicons, ontologies, thesauri or background knowledge base. This definition is deliberately broader and less specific than that in most of the standard references and intended to approximate common usage.
- *Coding system* – A terminology with attached identifiers or "codes".

We further point out the importance of distinguishing

- *Ontology* (Sensu philosophy): According to Quine[20], ontology is the study of what there is. Formal ontologies are theories that attempt to give precise formulations of the types of entities in reality, of their properties and of the relations between them
- *Ontology* (Sensu Informatics/computer science): First defined by Gruber as: "The conceptualisation of the entities in a domain"[21] (the ancestor to the more precise definition given above).

**18.** Alan Rector: Barriers, approaches and research priorities for integrating biomedical ontologies SemanticHEALTH D6.1, EC-supported FP6 SA, December 2007, www.semantichealth.org

**19.** Different authors refer to the meanings as "concepts", "universals," "categories". Note that the word "ontology" was borrowed from philosophy, and that there remain controversies concerning the extent to which the symbolic models referred to as ontolo-gies used in information systems should conform to principles laid down by philosophers for ontologies understood as part of the philosophical study of being.

**20.** Quine, WvO. On What There Is. Review of Metaphysics. (1948)

The difference is important, because the term "ontology" was "borrowed" from philosophy, but the two meanings are quite different. In philosophy, ontology strives to describe entities of a domain by their generic properties, whereas informatics and computer science focus on studying what is to be represented – and by extension the means of representation. The test of a computer science ontology is whether or not it is useful in information systems. This may, or may not, correspond to what any given school of philosophy considers to exist "in reality".

There is continuing controversy concerning the appropriate use of principles from philosophical ontology in the study of ontologies in informatics. On the one hand, insights from the philosophical study of ontologies often throw light on difficult issues. On the other hand, the distinctions and restrictions advocated by philosophical ontologists sometimes seem at best irrelevant and on occasion actively counterproductive to the use of ontologies in information systems. (The most vociferous advocate of the philosophical approach is Smith[22]). However, the philosophical approach to ontology may lead to a more principled and precise delineation of the meaning of basic categories or relations (e.g., in the OBO relation ontology[23] and BioTop[24]).

### 4.2.2 THE PROMISE AND PROBLEMS OF ONTOLOGIES

There are very few scientifically noteworthy items of knowledge that are truly "ontological" in a strict sense, whereas really "interesting" pieces of biomedical knowledge cannot be expressed by formal ontologies in a straightforward way[25]. Nevertheless, semantic interoperability can and should be addressed by the application of shared clinical terminologies and ontologies.

The primary goal of ontologies and terminologies is to enable the faithful exchange of meaning between machines and between machines and people and not to represent the state of the art of domain knowledge. However, the exchange of meaning is, in many cases, not sufficient for achieving SIOp. The electronic health record registers (and clinical information models) represent not only "what is", i.e. the reality of the patient, but also "what is known", i.e. the epistemic state of the documenting health professional[26]. Patients cannot be classified solely on their patho-physiological state, but must often also be classified in terms of what is known about them. This is reflected by codes such as "infection of unknown origin". Here is the key distinction between ontologies and information models: whereas ontologies represent what is always true about the entities of a domain (whether or not it is known to the person that reports), information models (or data structures) represent the artefacts in which information is recorded. Such information may be incomplete and error-laden which needs to be accounted for in the information model rather than in the ontology itself.

Additionally, practical experiences with biomedical ontologies[27],[28] have shown that important assertions such as "smoking increases the cardiovascular risk" (which is defined in terms of population-specific background knowledge) are of utmost importance for specific retrieval requirements of ontology users although they are difficult or nearly impossible to express by the representational formalisms used in biomedical ontologies.

Assuring the interplay of ontologies, terminologies, information models, and background knowledge bases with clearly defined tasks and interfaces is therefore a major desideratum[29].

### 4.2.3 SNOMED CT AND IHTSDO: CURRENT STATUS AND PROSPECTS

Since the International Health Terminology Standards Development Organisation (IHTSDO) set off, on a global level, to promote SNOMED CT as a resource devised to enhance global health by facilitating better health information management, any discussion of medical terminologies and ontologies is obliged to objectively assess the current status of this resource, to evaluate its fitness and to anticipate its impact on the health informatics scenario in the upcoming decades. First of all, the emergence of a kind of worldwide standard terminology should be appreciated. The fact that SNOMED CT has not been truly open was a major barrier to large-scale international development. However, things have improved since the creation of the IHTSDO: SNOMED CT research and evaluation licences are free of charge; the technical specifications are open, as is the collaborative website, and everybody can participate in working groups.

SNOMED CT's structure and implementation remain considerably flawed in many aspects.[30],[31] Although the

**21.** Gruber TR. Toward Principles for the Design of Ontologies Used for Knowledge Sharing. Journal of Human-Computer Studies. 1993;43:907-928.

**22.** http://ontology.buffalo.edu/smith/

**23.** Smith, B., Ceusters, W., Klagges, B., Köhler, J., Kumar, A., Lomax J., Mungall C., Neuhaus F., Rector A.L., Rosse C. (2005). Relations in Biomedical Ontologies, Genome Biology. 2005; 6 (5).

**24.** Schulz, S., Beißwanger, E., Hahn, U., Wermter, J., Stenzhorn, H. and Kumar, A. (2006). From GENIA to BioTop – Towards a Top – level Ontology for Biology. 4th International Conference on Formal Ontology in Information Systems (FOIS 2006), Balti-more, USA, November 2006, 103 – 114.

**25.** Schulz S and Jansen L (2008).. Molecular Interactions: On the Ambiguity of Ordinary Statements in Biomedical Literature. Applied Ontology.

**26.** Bodenreider O, Smith B, Burgun, A.: The Ontology-Epistemology Divide: A Case Study in Medical Terminology, Proc. FOIS-2006, Torino, Italy (2004)

**27.** Advancing Clinico-Genomic Trials on Cancer: project website (January 2006) http://www.eu-acgt.org

**28.** Integrated Biomedical Informatics for the Management of Cerebral Aneurysms (@neurIST): project website (January 2006) http://www.aneurist.org

**29.** Stenzhorn H, Schulz S, Boeker M, Smith, B.. Adapting Clinical Ontologies in Real-World Environments. Journal of Universal Computer Science, 2008

**30.** Schulz S, Suntisrivaraporn B, Baader F; SNOMED CT's Problem List: Ontologists' and logicisnas' therapy suggestions. 2007; Medinfo 2007: IOS Press; 802-806.

**31.** Rector A, Brandt S, Kola J. (2008) Why do it the hard way? The Case for an Expressive Description Logic for SNOMED. Pro-ceedings of the KR-MED 2008.

use of description logics[32,33] is advertised as an unparalleled advantage of this terminology, there is still no consensus in the SNOMED CT community about the domain SNOMED CT actually represents. Whether the concept "Chest Pain" is instantiated by the pain in my chest or by my physician's record entry (referring to my chest pain complaint) or by both, remains unclear. Note that description logics concepts can only be adequately interpreted (and unanimously modelled) if there is a consensus to which kind of entities they extend. Another issue not thoroughly addressed by SNOMED is the fact that its rather inexpressive variant of description logics (justified by scalability requirements) restricts axioms to very simple patterns that do not fit all purposes. This leads to inadequate statements especially with regard to negated contexts. Other shortcomings result from an idiosyncratic way of representing compositional hierarchies in anatomy. The consequences sometimes can even easily be spotted by laypersons, e.g. that "Amputation of Toe" is a kind of "Amputation of the Foot". (An encouraging fact is that a solution to this problem has been drafted in cooperation with SNOMED terminologists).[34]

For the time being we must therefore state that SNOMED CT is only fit for purpose as a controlled vocabulary and system of managed identifiers, but it can not yet safely be used as a source of semantics, which precludes its reliable binding to EHRs and the reliable use of post-coordination. Certainly, its scale is a major barrier to progress. It currently consists of over 400,000 concepts but there is evidence that only a small fraction of this material has ever been used and that the effort invested into SNOMED CT subsets seems to absorb more effort than the maintenance of the central terminology. The large scale and poor semantic expressiveness of the SNOMED CT relations and hierarchies considerably diminishes their usefulness. An equally severe problem is the nearly total lack of natural language definitions that are a major requirement to disambiguate entities clearly, to explain meanings to the uninitiated, or to find the correct translation.

Evidence-based results of SNOMED CT's fitness for purpose are missing. Yet the IHTSDO is aware of most of the SNOMED CT's shortcomings and is undertaking a major redesign effort for the subdomains of substances, organisms, observables, anatomy, events, conditions, episodes. A positive signal is the effort to align these SNOMED CT with existing terminologies, especially the Foundational Model of Anatomy (FMA). As anatomy constitutes a pivot discipline for all medicine, a common anatomical ontological reference for different terminological projects (SNOMED CT, ICD-10) would constitute an important advance.

### 4.2.4 GENERAL CONSIDERATIONS ON THE ROAD TOWARDS SEMANTIC INTEROPERABILITY

SemanticHEALTH has identified various issues on the road to enable semantic interoperability by the use of high-quality, new generation terminologies and ontologies. From a more technical perspective, it is recommended to focus the terminology development on concrete, immediate needs and real use cases with expected high benefits and low costs. Terminologies should have a well defined scope and purpose and be delivered against well defined, realistic time scales. It is recommended to separate ontology, language and interface, make it multilingual and multicultural, and to focus on quality assurance and reproducibility.

From an organisational perspective, all actions must aim to be imbedded in or aligned with long term institutions that can be sustained, healthcare providers and systems vendors must be involved effectively (furthermore, engagement of vendor and provider groups in any effort should be a prerequisite for funding), the terminologies must be 'owned' by their key end users, the terminology development must become coordinated with EHR and decision support developments and address multilingual and cross-cultural issues.

The desirable outcomes can be divided into content (what is actually in the terminologies and resources), tools and technology (what software is needed to make it possible to work with them) and process (what is required to sustain them). In terms of content, a focus of efforts should be on the user-centred development of a set of interoperable, scalable, modularisable, flexible, and adaptable ontology-based, and computable biomedical terminologies. They should include unambiguous formal and natural language definitions for all terms; a clear statement of the scope, purpose, formal foundations and technical specifications of terminology systems; lexical and linguistic support in all important European languages. Further important requirements are the availability of knowledge representation resources and formalisms using the terminologies, sufficient to support the needs of patient care, public health, clinical research, and health service management. In addition comprehensive maps to major classification systems; availability of a library of adaptable information models along with bindings to terminologies/ ontologies; multilingual language generation based

**32.** Spackman KA, Campbell KE. Compositional concept representation using SNOMED: towards further convergence of clinical terminologies. Proc AMIA Symp. 1998:740-4.

**33.** Baader F, Calvanese D, McGuinness DL, Nardi D, Patel-Schneider PF, editors (2003). The Description Logic Handbook. Theory, Implementation, and Applications. Cambridge, U.K.: Cambridge University Press.

**34.** Suntisrivaraporn B, Baader F, Schulz S, Spackman KA: Replacing SEP-Triplets in SNOMED CT Using Tractable Description Logic Operators. Artificial Intelligence in Medicine Europe 2007: 287-291

**35.** See SemanticHEALTH Deliverable D6.1, see esp. Annex 1, available for download at www.semantichealth.org

**36.** http://www.nlm.nih.gov/research/umls/Snomed/snomed_problem_list.html

**37.** For instance, a combination of the capabilities of the SNOB browser growing out of GALEN and Protégé-OWL environment linked to UK projects and the US National Center for BioOntologies should be encouraged.

on the above; as well as multilingual automatic encoding from text to support the above (to be coupled with voice recognition) are further important requirements.

### 4.2.5 RECOMMENDED ACTIONS

The selected recommendations address actions focusing on content, tools and processes in the development of terminologies. In addition, these actions are prioritised in time and respective implementation level needs.

***a) Areas needing adoption or short term action***

The following areas require actions centred on the content of clinical terminologies:

- A careful, methodologically sound, unbiased and public evaluation by independent evaluators of what SNOMED CT, in its current state, can and cannot be used for safely.
- Demonstration of a semantically sound and well quality assured reformulation of one or more suitable subsets of SNOMED CT[35], in order to provide evidence for long term decisions on the role of SNOMED CT in Europe. The timescale for this effort is seen as not more than three years. IHTSDO member states should take the lead. The effort should focus on a particular application – e.g., sensitivities and allergies – and a manageable and representative size (e.g. not greater than 25,000 concepts). The VA/Kaiser Permanente subset[36] or, alternatively a subset relevant to the ICD-11 effort might be chosen. Given the evidence from this effort, the remaining states, and the EC itself, would be in a position to make informed decisions about how to interact with SNOMED CT in the future. The criteria for success should be that using the reformulated version to create subsets is a) less resource intensive and b) more repeatable and reliable than existing methods.
- Open collaborative development of ICD-11 using Web-based technologies, which could serve as a demonstration for other open collaborative developments. Where possible this should be mapped to the semantically sound subset of SNOMED CT, or if necessary, to SNOMED CT itself.
- Harmonisation and cross mapping of major terminologies, including LOINC, DICOM, ICD 10/11, ICPC. Starting with the work on ICD-11, there should be a major effort to bring greater convergence and harmonisation of the various national terminologies. As for all these terminologies, human anatomy is a common point of reference (most diseases, signs, symptoms, procedures can only be exactly described and defined referring to anatomical entities), it would be a major achievement and a solid basis for cross mapping if those terminologies referred to the same source of anatomical terminology, such as, e.g. provided by (a subset of) the Foundational Model of Anatomy.
- Adoption of a clear policy that endorses the mapping of all terminologies mapped to UMLS CUIs and LUIs, either by their originators or in collaboration with the US National Library of Medicine.
- In the context of Translational Medicine, convergence on a common terminology for clinical trials and longitudinal studies should be enhanced, such as the creation of a common terminology for BioBanking.

The following areas require actions focusing on tools and technologies:

- Development of free authoring environments and toolsets to support terminologies and ontologies. These should be capable of handling compositional terminologies in general and SNOMED CT in particular.[37]
- Development of a prototype network of terminology and archetype servers for European countries. The development of tools for coordinated authoring of terminologies and archetypes should lead, in the medium term, to a network of "just in time"' centres/web sites where users can get quick responses to their needs.
- Linking of tools to developments in Web 2.0, Social Computing, the Semantic Web, Text Mining, and related disciplines in order to implement collaborative web-based workflows. The development of the ICD-11 within a collaborative framework could be an ideal vehicle for this task.
- Development of SNOMED CT browsers and definition of a set of core browsing features which are defined and harmonised across tools to allow for a standard browsing experience across environments.
- Development of environments for coordinated development of terminologies and medical record standards, starting with the addition of facilities for linking terminologies to editors for the Archetype standard and CEN EN 13606. This is a high priority task.

The following areas require the implementation of specific processes, such as

- Establishment of a European centre of expertise for collaboration with the National Cancer Institute on the use of its ontology and terminology resources, with the goal that groups should be able to build on this platform with only modest cost.
- Elaboration of priorities for research on combined approaches to knowledge representation, ontologies, and web technologies investigating how to achieve the optimal balance between the various technologies for clinical applications.
- Widespread engagement with both vendors and providers on requirements, gaps, and irrelevancies. Engagement of vendor and provider groups in any effort should be a prerequisite for further funding.
- Involvement of the prospective primary user community to an extent that they feel they are the 'owners' and 'custodians' of this resource.
- Evidence on several important questions should be achieved such as on SNOMED CT implementation and migration cost, interrater reliability, accuracy and patient safety benefits / hazards.
- Formal links and a European centre of expertise for collaboration with the US National Center for BioMedical Ontologies (NCBO) and collaboration with them on their tools and kits and with the National Cancer Institute on its CaBig and EVS systems.

- Support for open tools for terminologies that link up to SNOMED CT
- Development of language technologies: text extraction to build new terms and encode natural language; text generation to present and for quality assurance (QA).

***b) Areas* needing widescale evaluations (medium term)**

- The use of SNOMED CT with Archetypes and HL7 v3
- The use of SNOMED CT to map to ICD 10/11
- Use of Social Computing mechanisms to QA and provide feedback on ontologies and terminologies, as well as to assist with translation of terminologies
- Identification and cataloguing of the terminologies actually used in the various European states, including, and in particular, the terminologies for drugs, adverse reactions, and sensitivities, as a preliminary for harmonisation.

*Areas needing investment (medium term)*

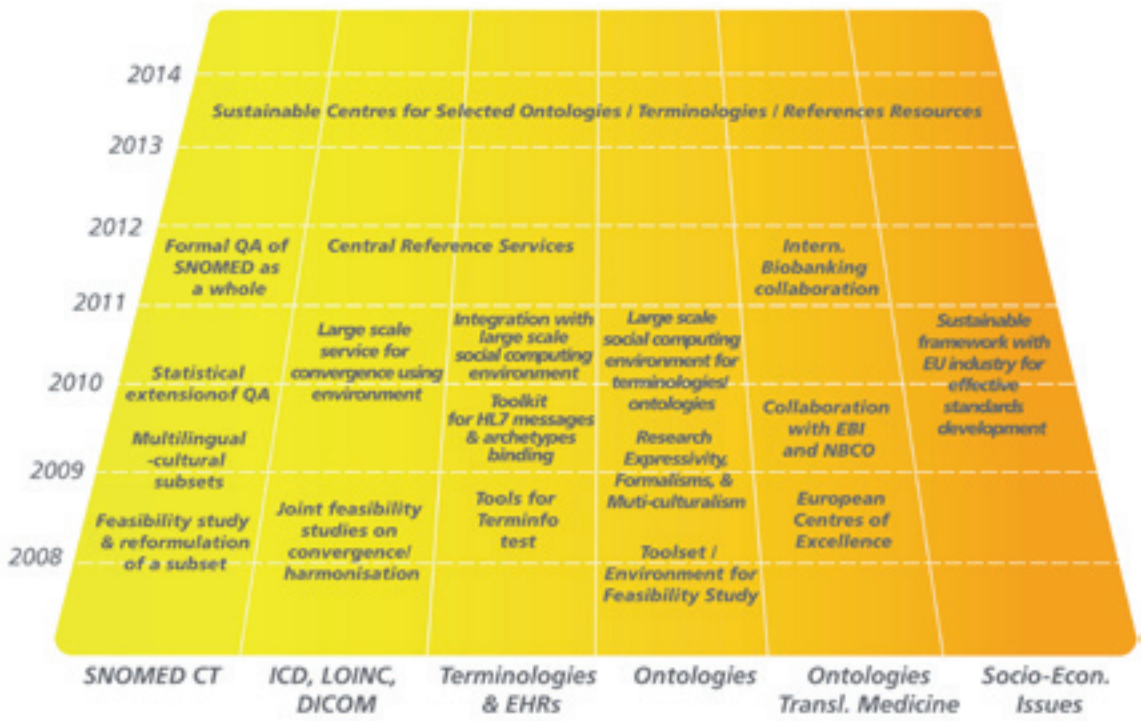
- Establishment or support for a major network of tool providers and terminology service providers for Europe
- Development of human capacity and skills in terminology and ontologies
- Development of methodologies and tools for binding of archetypes and HL7 v3 messages to terminologies and ontologies
- Engagement with the IHTSDO and HL7 with sufficient focused resources to have a major impact.

***c) Areas needing further (focused) research (long term)***

- Ontology driven architectures for clinical medicine
- Optimal balance between "Google-like" and semantic techniques for interoperability
- Evaluation metrics for ontologies and terminologies for use in Healthcare IT, including the relevance of "good ontological practice" to practical software success.

The summary roadmap for terminologies and ontologies is reproduced in Figure 2:

FIGURE 2: SUMMARY ROADMAP FOR TERMINOLOGIES AND ONTOLOGIES



The key long-term goal in the domain of bio-medical and clinical research, and translational medicine is to achieve a step-change in the speed and effectiveness of clinical, experimental and pharmacogenomic research whilst respecting patients' rights to privacy and informed consent. The following use cases have been identified for further investigation: multicentre studies and trials; repositories, biobanking; personalised medicine based on genetic and genomic analyses.

Key items of an action plan to realise the ontology roadmap specifying in some detail issues of content, process and tools are provided in Table 1:

**TABLE 1: ONTOLOGY ACTION PLAN**

| YEAR | CONTENT | PROCESS | TOOLS |
|---|---|---|---|
| 2008 | • Open ICD 11 with SNOMED CT Mappings<br>• Reformulation of SNOMED CT | • Establish open collabora-tive framework for ICD 11<br>• Begin to establish mechanisms for industrial | • Open tools for ontology development<br>• Open Web 2.0 tools to support ICD 11 |
| 2009 | • Quality assurance metrics for SNOMED CT fragment<br>• First translations of SNOMED CT fragment | • Extend industrial in-volvement | • Develop tools for linking and binding terminologies and archetypes<br>• Establish open social site for clinical terminology |
| 2010 | • Continue | • Establish mechanisms for reformulation of SNOMED CT fragment | • Extend and test open tools for terminology and arche-types, possibly incl. HL7 |
| 2011 | • Reassess and create long term plan for selected ter-minologies including limits on scope | | • Establish European Net-work of Terminology Servers |
| 2012 | • Review and reassess interoperability | | |

In summary, the key milestones on the road to SIOp include: a semantically sound SNOMED CT fragment supported by tools and organisation; a social/collaboratively built ICD-11 with widespread support; the mapping of ICD-11 to a semantically sound subset of SNOMED CT; the establishment of a set of widely used Web-based terminology services for access to, quality assurance, and feedback on clinical ontologies.

# 4.3 Public health

One of the greatest 'added-values' of digitalisation of individual health information is enabling their combination, aggregation and analysis at population level. This will allow to compute various indicators, benchmarks and trends of public health issues with respect to:
a) populations, groups
b) settings, facilities
c) regions, geographic units, and/or
d) environmental variables.

The digital public health and epidemiology vision of the SemanticHEALTH project foresees the integration of anonymous data from individual EHR records into interoperable data systems with full coverage across populations at facility, regional, national and international levels.[38] The following areas have been recommended for further investigation:

- Facilitating international statistics
- Assessment and surveillance of outcomes and diseases
- Improving patient safety
- Underpining population health research

The World Health Organization (WHO), partner in SemanticHEALTH in charge of public health issues, launched a "Public Health Informatics Key Informant Survey" (PHIKIS) in order to evaluate the requirements and the feasibility and prioritisation of various use cases for public health informatics[39]. The priority use cases for digital public health as identified through this survey are the :

- digitalisation of mortality statistics
- use of routine laboratory data
- use of ePrescriptions.

This ranking reflects concerns for feasibility and utility. The top three use cases according to the priority attached by respondents are mortality statistics, administrative costing purposes and public health reporting. Further applications could cover adverse drug reaction monitoring, disease reporting according to international health regulations, aggregation statistics for patient flows and calculation of case-mix groupings using ICD, ICF and grouping algorithms from diagnosis related groupings (DRGs).

The consistent use of electronic health / person-based health records and the interconnecting of health service provider systems are key prerequisites for this vision to become reality. Furthermore, research is needed to prove the comparability of digital information with analogue traditional measures used for monitoring and evaluation in various health information systems. Finally, a standardised approach for representing and sharing of public health indicators (e.g. rates for mortality, morbidity or vaccinations to be expressed as public health archetypes) needs to be developed and implemented.

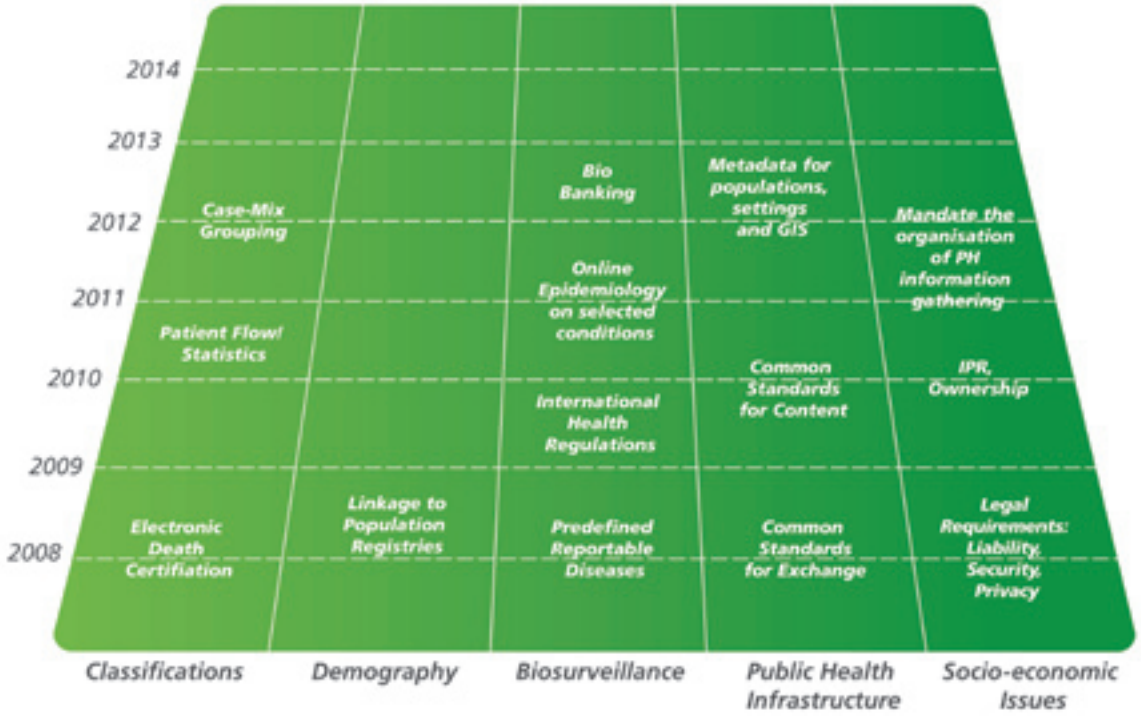
**38.** Bedirhan Üstun: Public Health Use Cases and Health Information Standards, SemanticHEALTH D5.2, 2008

**39.** http://www.who.int/classifications/phi-kis/en/

Recommended actions in the field of public health therefore include the development of common standards that will allow data-exchange of preidentified variables from an individual EHR, and compilation and comparison of that data across regions, time and populations.

The full set of recommended actions is summarised in Figure 3:

FIGURE 3: SUMMARY ROADMAP FOR PUBLIC HEALTH




Further action should aim at the establishment of:

- National Centres for multilingual, multicultural adaptation of international classifications and terminologies, including SNOMED CT, linked in a well-managed European Network of Competence Centres, to be expanded globally
- A European and global Network of Terminology Servers.

Sustainability and scalability need to be assured and the terminology servers need to be maintained in order to be useful.

Finally, there is a need for action on the legal framework providing the enabling environment for public health action. There needs to be clarity on security and privacy regulations of public health data, clarity on mandate and responsibility of those actually carrying out the data aggregation, and finally rules on liability, should damage to individuals arise in the process.

# 4.4 Socio-economic issues

### 4.4.1 BENEFITS AND COSTS OF SEMANTIC INTEROPERABILITY

Reaching high levels of interoperability is a resource-intensive task. The general economic challenge of investing in ICT in the health domain is to maximise the benefits from eHealth, given the constraints in resources. Interoperability plays a significant role in this optimisation equation. It is often essential in realising the benefits from eHealth investments, but it may also consume a significant share of the available resources.

Issues and considerations related to costs concern implementation and utilisation, whereas utilisation is also critical for the realisation of benefits. Among others, some of the cost factors related to semantic interoperability include the development, translation and maintenance of terminologies, change management requiring additional training and education, and the harmonisation of data collections.

The following questions arise: (1) who is responsible for development and implementation, (2) who will pay, and (3) who will accrue the benefits. The costs of healthcare IT programmes now exceed ten billion Euros per annum across the Union. The health policy system cannot currently associate the benefits of interoperability with those who must pay for it. Therefore a convincing demonstration of the benefit of migration from legacy to interoperable systems is required in order to justify public intervention.

Diversity of healthcare systems, language, clinical speciality, maturity of economic development, and IT acceptance cause severe market segmentation. The total cost of lack of standard terminology may be very large as well as the costs of the creation of ad hoc terminologies. However, that spending can rarely be aggregated for accounting purposes. Net savings are therefore difficult to measure, whereas the cost of a major modification to any central standardised system are highly visible. The commitment of large vendors to interoperable solutions (except HL7v2, LOINC, DICOM) has been limited.

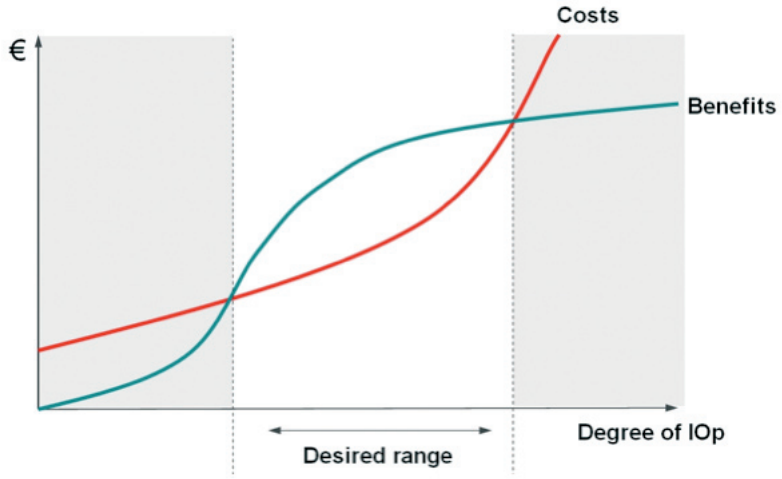
But even though the improvement of SIOp comes at a cost, the expected benefits are also considerable. These relate mainly to the speed and consistency of accessing meaningful health related data. Key aspects pertain to the medical staff saving work time, gaining efficiency and improving safety and clinical outcomes through better access to patient information across disciplines, care settings and even countries. The potential for efficiency gains has been highlighted by studies of the RAND Corporation or the Centre for Information Technology Leadership (CITL).[40] In the same vein, when SIOp aims at a doctor-patient relationship, the role of the patient is strengthened and health services may be improved. SIOp based collaboration can also lead to a decrease in reaction time to global threats such as pandemics. From a public health perspective, further benefits could be derived from being able to use richer clinical detail, leading to improvement and greater confidence in information used for audit, planning, and performance management. A prospective analysis of costs and benefits for different degrees of SIOP and different purposes suggests the following most promising areas for shorter-term investment. For instance, the achievement of partial SIOp for a minimal data set is expected to yield high added value at moderate cost. A high degree of SIOp of EHRs for direct patient care purposes would be very costly, but could also yield very significant benefits. For the purposes of public health research, striving for partial SIOp should come at moderate costs and yield high benefits. Finally, a high degree of SIOp for the purpose of research and knowledge sharing would yield very significant benefits, albeit at high costs.

The benefit cost dynamics of striving for full semantic interoperability become more apparent when understood as a probably nonlinear phenomenon. SIOp is not a binary variable, but rather a scale reaching from zero to full IOp. Various levels will imply different benefits and costs, and therefore it will be of critical importance to better understand and estimate these relationships to determine optimal levels of IOp. This hypothetical trade-off is illustrated in Figure 4. Underlying assumptions - to be tested by empirical evidence and for different situations - are that after initial set-up costs establishing higher degrees of SIOp will involve costs which increases in a nonlinear fashion as higher degrees of SIOp are desired. On the other hand, initial benefits from low degrees of SIOp are assumed to be low, to increase fast after an initial threshold is overcome, and to level out the more one approaches full SIOp. Only evidence can tell us whether there exists indeed an optimal range for the degree of SIOp, and where this range may be.

**40.** The key messages from the RAND impact study are that ICT for Health adoption may become associated with large potential benefits: Girosi, Federico, Robin Meili, and Richard Scoville (2005). Extrapolating Evidence of Health Information Technology Savings and Costs. Santa Monica, CA.: RAND Corporation.. The CITL study considered that standardised information exchange systems could result in net savings of as much as 5% of current US healthcare expenditure. Walker J, Pan E, Johnston D et al. (2005). The Value of Health Care Information Exchange and Interoperability, Health Affairs online edition. See however the critical methodological review of these and other studies in: The Congress of the United States - Congressional Budget Office (2008). Evidence on the Costs and Benefits of Health Information Technology, May.

**41.** The net (societal) benefit of interoperable Electronic Health Record Systems in different settings has been studied from a bene-fit-cost perspective in the European Commission funded study EHR IMPACT, www.ehr-impact.eu

FIGURE 4: BENEFITS AND COSTS OF SEMANTIC INTEROPERABILITY DEGREES




The main recommendations stemming from the analysis of the socio-economic aspects of SIOp are:

- Researchers need to investigate and analyse the approximate shapes and behaviour of the benefit and cost curves of SIOp in different settings, i.e. undertake well-grounded Benefit-Cost-Analyses (BCA).[41]
- Technical and organisational developments should be guided by the results of prospective BCA, and in particular the estimation of an optimal degree of SIOp. Regular impact assessments would help to evaluate progress and adjust strategies if necessary.
- To provide for and ensure the sustainability of given SIOp developments, continuous cooperation among stakeholders, and thus realisation of (societal) benefits over time together with the individual (private) incentives have to be understood and if necessary adjusted.

SIOp comes at a cost, initially as well as continuously, to be covered by various stakeholders. A BCA approach, once validated, can be used to determine an optimal development path and for allocating the burden in an efficient way. Furthermore, it can help to compare the relative benefits of institutional solutions like single (centralised) European or national centres vs. (decentralised) collaborating centres, the impact of European cooperation, of public versus private ownership. It is also recommended to create reference sites within, e.g., hospitals across various types of healthcare systems in the EU where the clinical, financial and operation improvements due to the usage of terminologies, ontologies and electronic health record systems are documented and visible to visitors.

Looking further into the future, it is important to understand and take into consideration the workflow and daily routine of medical personnel and to suggest incentives for a successful and sustainable deployment of the roadmap. It is recommended to undertake a thorough investigation of incentives for development, implementation, maintenance, and utilisation of SIOp based collaborative practices.

### 4.4.2 OTHER HORIZONTAL ISSUES

A large number of further horizontal, non-technological issues which impact on the cost and benefit structure of semantic interoperability and on the organisational and political space, in which such efforts are to be implemented, need an indepth analysis.[42] The political objectives of Member States may generate different priorities for semantic interoperability. Further, key aspects of the legal and regulatory environment within which semantic issues are embedded play a significant role. In various countries, the application of certain coding schemes like ICD may be required by law, and also infrastructure institutions necessary to sustain such systems may be part of the public healthcare system of a Member State. These have to be considered when planning to improve semantic interoperability.

Next to the political and regulatory issues, organisational and even cultural issues can have an important impact on the trajectory and effectiveness of efforts towards semantic interoperability.[43] Examples include changing care pathways, increased collaborative working and exchange of information between providers. Stakeholder involvement is a key factor to be considered. Increasing awareness of societal benefits and motivating stakeholders to take an active part in relevant processes and organisations are crucial success factors. Last but not least, the multicultural and multilingual aspects of health information exchange play a key role in a European dimension.

**42.** For detailed analysis see D3.1 (Socio-economic issues), D7.2 (Final Roadmap), at www.semantichealth.org

**43.** Two examples of this wide literature include: Coiera E (1999) The impact of culture on technology: How do we create a clinical culture of innovation? Medical Journal of Australia 171:508-9. Also: Narine L and Persaud DD (2003) Gaining and maintaining commitment to large-scale change in healthcare organizations, Health Services Management Research 16:179-187.

### 4.4.3 SOME EVIDENCE ON COSTS

Whereas estimates on aggregate benefits from terminologies at the health system level are missing, some overall cost estimates are available. In the 1990s, the United Kingdom National Health Service (NHS) purchased the so-called Read Codes[44] and further developed them. It involved over 2000 clinical professionals in the development of Version 3, at a cost estimated at that time of up to £30 million ($45m).[45]

According to IHTSDO, the organisation governing the SNOMED terminology, the cost of de-veloping a comprehensive clinical terminology is estimated to be between $US 25m and $US 50m. The annual maintenance cost is expected to be more than $US 8m per year – which is the estimated cost for the IHTSDO to maintain and improve SNOMED CT as a global re-source. These estimates are based on the historic costs borne by the College of American Pathologists and the UK National Health Service.[46]

In Sweden, which decided to join the SNOMED consortium in 2007, the National Board of Health and Welfare was commissioned by the central government to submit in 2006 a prelimi-nary assessment of what the introduction of a national terminology system (SNOMED CT) into the health system would entail in terms of demands on the National Board of Health and Wel-fare and at what cost. As a general observation, the Board noted that this would involve a long-term financial undertaking for all affected stakeholders such as the state and county councils/municipal authorities as well as private care providers. Dental care and social serv-ices were to become involved, too, and a particular focus on dealing with interdisciplinary ter-minology was planned.

Envisaged application areas for SNOMED CT in Sweden include health information, medical advice and services as well as medical records systems, administrative support systems, ePrescribing support and a national patient summary. The total cost for the project – including project management, translation into Swedish, further development and maintenance, support tools, training and liaison – were at that point in time estimated at about € 14m for the four year introductory period 2007 to 2010, and € 1.7m per year thereafter.[47]

To further clarify cost issues and complement them with well founded benefit estimates, it maybe worthwhile to focus initially on a more limited application (as suggested above in Section 4.2), like a small scale pilot within a hospital or region, using small subsets of SNOMED CT and ICD 10 (e.g., limited to sensitivities and allergies). Another pilot could include a linkage between electronic patient record data and mortality statistics for public health purposes (death certificates for a limited number of leading causes of death).

Attempts to develop a cost estimation model to predict the efforts involved in building, reusing and maintaining ontologies in information systems have been undertaken within the Semantic Web community.[48] Various factors related to the scope and depth of ontology building, the complexity of the domain and availability of reusable material impact on the cost estimates. Efforts for maintenance, translation, tooling, integration into information systems, training, evaluation, quality assurance, and assessment of understandability by humans are other key elements. This research has indicated that short term action in ontology building should focus on mobilising adequate expertise and encouraging distributed development.

**44.** They were later merged with SNOMED RT to become SNOMED CT.

**45.** National Audit Office: NHS Executive: The Purchase of the Read Codes and the Management of the NHS Centre for Coding and Classification. London, HMSO, 1998.

**46.** http://www.ihtsdo.org/about-ihtsdo/faq/ [accessed July 2008]

**47.** Swedish National Board of Health and Welfare (2006). SNOMED CT: Project Preliminary assessment of measures and costs, 27-11-2006, Stockholm.

**48.** Elena Paslaru Bontas, Malgorzata Mochol (2006): Ontology Engineering Cost Estimation with ONTOCOM. KnowledgeWeb Network of Excellence, Technical Report B 06-01, March 24, 2006.

# 5 Summary and outlook

## 5.1 Recommendations

The SemanticHEALTH roadmap has pointed to various challenges and the respective domains where action is required on the path to achieving semantic interoperability in support of European health services. A policy of incremental steps and a focused, modest approach to terminology development in an open, collaborative environment is the ultimate recommendation following from the project's work.

A strong and coordinated effort is recommended to effectively engage with the relevant stakeholders policy makers and health authorities, service providers, health professionals and professional organisations, citizens, academic organisations and research funding agencies, standards development organisations, industry - primarily healthcare ICT, insurance sector, and especially the national competence centres / national initiatives in charge of EHR / eHealth systems implementation. Further action should aim at the establishment of sustainable national bodies (e.g., national centres for multilingual, multicultural adaptation of international classifications and terminologies), linked in respective European networks.

In order to demonstrate the impact of semantic interoperability it is also recommended to create reference sites within hospitals across the EU where the clinical, financial and operation improvements due to the usage of terminologies, ontologies and electronic health records are documented and visible to the health community. Key performance indicators need to be identified with regard to the impact of terminologies, ontologies and other semantic tools on patient safety, quality of care, efficiency etc. A comparison should be made between the outcomes before and after the implementation of semantic technologies. Costs of implementation and maintenance should be considered. The sites could also provide training related to SIOp and outcomes.

An overview of key recommendations is provided in Table 2. Table 3 on the following page illustrates the interdependencies between milestones in each domain.

In conclusion, it is essential to invest in the coordinated production of tools supporting the development and deployment of terminologies and archetypes, recognising the distinction between tools that support their development processes and those that support the utilisation of the products. EU instruments should include R&D projects (via the Framework Programme) and Networks of Excellence. Short term goals should include the authoring of archetypes, the development of ontology-driven multi-lingual tools, and the binding of terminologies to archetypes and information models. These should be complemented in the medium term with the development of the tools referred to above, including archetype based tools for binding information models with terminologies. Another essential task is the creation of semantically sound and focused subsets of SNOMED CT and ICD 10 that have immediate relevance to the health improvement priorities of Member States.

## 5.2 Outlook

For more than 15 years, the European Commission has recognised the importance of terminologies and interoperability by funding research in these fields. In this publication, we have identified key actions to achieve faster progress towards a more consistent representation of clinical meaning across European and global health systems. Concrete initiatives by Member States, their competent centres, healthcare providers, industry and other stakeholders together with the research community are now needed to realise the semantic interoperability vision to indeed reap the benefits from a wider implementation of eHealth solutions.

Semantic interoperability is a local as well as a global issue, and is taken forward by many national and international efforts. Based on the foundations laid by recent research, Europe has a unique window of opportunity to make faster progress in this area because many of its health system stakeholders are very well aware of the cultural and linguistic challenges the new model of collaborative healthcare involves, and also of the opportunities arising to meet the growing needs for the sustainability of our health systems. Most Member States are presently investing in eHealth strategies and infrastructures, and the European Commission can be expected to provide continuing help and coordinating mechanisms to support them. We hope that the analysis and recommendations presented above will help enable future eHealth investments to deliver early, tangible and valued outcomes on the journey towards semantic interoperability.

TABLE 2: OVERVIEW OF KEY RECOMMENDATIONS

| | *Deployment*-oriented R&D recommendations | | *Research* focused R&D recommendations | |
|---|---|---|---|---|
| | **Short term** | **Mid-long term** | **Short-term** | **Mid-long term** |
| **Electronic Health Records** | Agree to use *archetypes* as the standardised approach to *representing and sharing clinical data structures* across the EU (large pilots).<br>Use the *13606* standard *to structure* the EHR / patient summary.<br>Support a *Network of Excellence* to facilitate the use of archetypes. | Develop good practice in *archetype design and terminology binding to them*, and their governance.<br>Develop a *coordinated network* of archetype repositories accessible online. | Work on archetype *indexing* and ontology binding to archetypes.<br>Develop archetype editors, open source tools and web based facilities for *training clinicians*. | SIOp solutions for *Personal Health Records* (PHR).<br>Linking EHR data to educational materials and clinical evidence, to enable consumer engagement.<br>Support health professional decision making and education, improve adherence to evidence based care. |
| **Ontologies & Terminologies** | Develop a well QA'ed *subset* of SNOMED CT terms for a particular application experimenting with alternative representations.<br>Use this exercise to *test methods of binding terminology to information models*<br>Establish a *Network of Excellence* on ontology to facilitate development, maintenance and use of international terminologies such as ICD and SNOMED CT in different languages. | Develop widely used *server / services for communal editing*, QA and deployment of multilingual – multi-cultural clinical terminologies in cooperation with UMLS.<br>Promote good practice guidelines in collaboration with the EHR recommendations above.<br>Establish a network of national reference centres in ontology to make available structured multilingual terminologies allowing safe exchanges of meaning. | Develop and demonstrate *QA criteria for ontologies* in use including tests for inter-rater reliability, appropriateness *criteria for decision support tasks*.<br>Develop *ontology driven multilingual open source tools* and web training facilities for clinicians, public health specialists and terminology developers. | Establish the theoretical foundations and practical tools for a *coordinated set of representations for clinical meaning, data structures, and rules* (This remains a long term "Grand Challenge" although it will have shorter term intermediate objectives).<br>Develop a coordinated network of ontology repositories accessible online. |
| **Terminologies relevant for Public Health** | Establish the right framework for *aggregation of EHR data to produce public health indicators and alerts* (large pilots, e.g. linking EHR data to mortality statistics. | Based on priority use cases (e.g. patient summaries), define minimum interoperability criteria for public health use<br>Develop *EHR based regional public health data bases*. | Develop one segment of ICD 10 with, e.g., Protégé-OWL and refine language generators for at least 3 languages.<br>Develop *aggregation logic tools* and web based training facilities for clinicians, and public health specialists. | Develop mappings of terminological resources using common methodologies.<br>Develop a coordinated network of regional public health data bases based on international standards and agreed methodologies, and accessible online in multiple languages. |

| | Tools / Content / Processes | 2008 | 2009 | 2010 | 2011 | 2012 | 2014 | 2015 |
|---|---|---|---|---|---|---|---|---|
| EHR | Reference Model | Generic model for EHR communication | | | | | | |
| | Archetypes | Standardised represention/ Sharing of clinical data structures | Best practice Archetype design | Authoring/validation tools | Quality Assurance and certification | Repositories | Terminology binding | Care pathways |
| | Terminology Systems | Key Use Cases | Policies on SNOMED CT-term coordination | SNOMED CT subsets | Terminology server/term browser for SNOMED CT | Business rules for term coordination | Consistency test HL7 Terminfo trial standard | SNOMED CT global experience test |
| | Technology/ Visualisation | | | | | EHR visualisation applications | Adaptable clinical applications | |
| | Socio Economic Issues | EHR/terminology user training | Improve internationalisation across HC paradigm and cultural differences | Agree on semantic interoperability goals for PHR | | | Link EHR data to educational material | Acceptance evaluation |
| Ontologies | SNOMED CT | Feasibility study and re-formulation of SNOMED CT subset | Multilingual / -cultural SNOMED CT subsets | | Statistical extension of QA | Formal QA of SNOMED CT as a whole | Policy in conjunction with MS on future use of SNOMED CT hierarchies and relations | |
| | LOINC, DICOM | Consultations on issues related to LOINC & DICOM | | Actions arising from consultations | | | | |
| | Terminologies and EHRs | Tools/methods for Terminfo guidelines test implementation | Toolkit for HL7 messages and Archetypes binding | Integration with large scale social computing environment | | | Central reference terminology services | |
| | Ontologies | Toolset for SN CT feasibility study Ontology driven multi-lingual tools | Research: expressivity, formal-isms, multi-culturalism | Ontology mapping research | | Large scale collaborative ontology environment | Sustainable Centres for selected ontologies | |
| | Genomics/ Translational Medicine | European Centres of Excellence | Collaboration with EBI and US NBCO | international bio-banking collaboration on terminologies | | | | |
| | Socio Economic Issues | Establish sustainable framework with European industry for effective standards and terminology/ontology development | | | | | | |
| Public Health | Classifications | Electronic death certification | | | Patient flow/ statistics | Case-Mix Grouping | | |
| | Demography | Linkage to population registries | | | | | | |
| | Biosurveillance | Predefined reportable diseases | | International health regulations | Online epidemiology on selected conditions | Bio-banking | | |
| | Public Health Infrastructure | Common standards for exchange | | Common standards for content | | | Metadata for populations, settings and GIS | |
| | Socio Economic Issues | Legal requirements on liability, security, privacy | | | Legal requirements on Intellectual Property, ownership | | Mandate the organisation of PH information gathering | |

European Commission

***Semantic Interoperability for Better Health and Safer Healthcare***
Deployment and Research Roadmap for Europe
Authors: Veli N. Stroetmann (Ed.), Dipak Kalra, Pierre Lewalle, Alan Rector, Jean M. Rodrigues, Karl A. Stroetmann, Gyorgy Surjan, Bedirhan Ustun, Martti Virtanen, Pieter E. Zanstra

Luxembourg: Office for Official Publications of the European Communities
2009 —32 pp. — 21 x 29.7 cm
ISBN-978-92-79-11139-6

KK-80-09-453-EN-C

# Semantic Interoperability

## for Better Health and Safer Healthcare

For further information, please contact:

**European Commission**
Information Society and Media Directorate-General
Directorate H - ICT adressing Societal Challenges
Unit H1 - ICT for Health
Phone: +32 2 296 41 94
Fax: +32 2 296 01 81
eHealth@ec.europa.eu
http://ec.europa.eu/information_society/ehealth

Design & Printed by OIB

ISBN 978-92-79-11139-6

9 789279 111396

European Commission
Information Society and Media