

Health terminologies: Criteria for decision making

eHealth Working Group Workshop Brussels
Wednesday March 8, 2006

Pieter E Zanstra

Radboud University Nijmegen Medical Center
Co-ordinator EU SemanticHEALTH Specific Support Action:
Semantic Interoperability Deployment and Research Roadmap
P.Zanstra.mi@xs4all.nl

1 Motivation for this workshop

In the fall of 2005 an initiative was launched to establish a global standardisation organisation for the further development and implementation of SNOMED. Invitations to join this organisation have been sent amongst others to the member states. The information pack that is distributed during the meeting has the complete texts of that invitation. The organisers of this meeting can inform Member States individually on the national contacts where the invitations have been sent to.

This meeting is intended to further inform eHWG members and/or other officials in your Ministries who have a (partial) responsibility for health terminologies. Participants should bear in mind that though terminology is important for the full semantic interoperability between systems, it is not the exclusive domain for eHealth. There are many other parties that have a legitimate interest in health data, or information derived from health data. As a consequence those other stakeholders should be included in the decision making at some point in time. In the fall of 2004, the CEN/ISSS eHealth Focus Group finished a study on 'Current and future standardization issues in the eHealth domain: Achieving interoperability' one of the recommendations says, The Member States, with the Commission, should:

- *ensure the Europe wide referencing and easy access to the content of existing health coding systems based on registration of such systems by the Eurorec Institute;*
- *support the international convergence towards a common framework for formal representation, and eventually the development and maintenance of a multilingual clinical reference terminology. This effort should build on existing efforts in formal representation as GALEN, FMA and SNOMED, and be carried out in liaison with the WHO Family of International Classifications*
- *make the targeted reference clinical terminology publicly available free of charge;*
- *support a common approach to link national classifications of procedures, to support cross-border reimbursement of health care.*

An imminent question today is to join the Snomed SDO or not. But is it the right question? As a matter of fact there are many pending questions to be answered before. How to decide? Is it the right solution to the (perceived) needs? Aren't simple coded lists sufficient for the medium term? There is tension between freedom of expression in natural language versus the need for control to ensure qualitatively good data. What makes sense to be described in standardised language? Is international co-operation helpful or is it hampering domestic developments. And if we cooperate what is the most appropriate organisational form. These and many more questions are not easily answered.

The objective of this workshop is for Member States to gain a broader understanding of what terminologies are useful for in general, and in the eHealth context in particular.
--

We will bring you into contact with the experts in the field. You can consult them for your domestic questions during the meeting and afterwards. We present short statements of needs and some solutions from four member states. There is an in depth explanation on SNOMED itself. Some first SNOMED experiences from a non English speaking country and organisational issues are addressed in the afternoon. We will discuss what needs to be done

in general both at the regional/ national and the European/ global level, and what SNOMED can contribute to this initiative in particular. At the end of the day there is an hour for a round table discussion. We have asked participants to raise topics for discussion. The intention is to focus on the top 5 topics. This round table will be prepared and supported by experts from the SemanticHEALTH and RIDE projects.

The title for the workshop is 'Criteria for decision making'. Actually there is no clear set of such criteria. The decision problem is complex and multi-faceted. This workshop is not about 'the' solution. There will not be a single answer for member states or the EU as a whole at the end of this day. We aim at raising awareness at the European Union level, and to clarify key issues and priorities. This workshop will as much as possible bring issues to the table that need attention in different fora. At this point it seems there is a need of at least one or two more sessions which drill down on specific topics.

This one is just a start.

2 History¹

The healthcare delivery process is an extremely information intensive industry. This is the case for direct patient care as well as for registration and billing activities. Because the healthcare process takes a substantial part of the national budget, the society demands to know a lot of that process. Where the financial industry have Euros, Dollars, etc. as a relatively simple common ground, health language, the 'currency' of the health care delivery process, is much more complex, and still far away from being harmonised. To achieve unambiguous recording of patient data across the sector, a good common understanding of the words used is very important. Many have expressed the need for terminology that can be understood by all either directly, or after being 'translated'.

Direct patient care requires rather detailed terminologies to closely describe the patient's condition. Most of present day classifications like ICD and procedure classifications have not been designed for documenting direct individual patient care. They were primarily designed to describe at a aggregate level and for purposes of counting. Clinicians do not consider the language of existing classifications as their own language. This results in the development of many local lists and adaptations of varying quality and compatibility.

Ideally for any patient case there ought to be a detailed recording of findings, diagnostic procedures with results, diagnoses, and the reasoning that has led to such conclusions. Also treatment and results, as well as any acts that have been performed as part of the treatment plan need recording. This must be done in an unambiguous language within a structure that makes it reusable and useful for any qualified healthcare professional at any time, at any place. In that sense there is in essence no difference between sending patient data from one ward to another or for transfer of duties at the end of the day. Current practice shows that relatively little is recorded in a standardised way. Secondary care has just barely started with the introduction of a real electronic health record. Quite often one sees free text from which it is always easy to reconstruct clinical argumentation. Some documents are even stored as scanned images, which render them totally useless for automatic processing. Laboratory data and drug prescriptions generally are reported in a more formal way. The same is the case for discharge diagnoses for national registries, and procedures for reimbursement purposes. In some specialties there are national or international groups that have agreed on a common registry for research purposes. Unfortunately such activities are mostly uncoordinated and not supported by skilled terminologists.

The purpose of the registration of data is to have a faithful record of what is done with or for the patient. This recording must be useable for multi-disciplinary treatment of the patient. Also the increase of part-time doctors dictates that more doctors of the same specialty must be able to communicate about the same patient via one single (virtual) record. Simultaneous or consecutive medical attention to the patient requires insight and understanding from all participants in the treatment. This is a very ambitious goal, but the change of healthcare

¹ The following pages contain slightly adapted extracts from the Report 'Introduction of clinical terminology in the Netherlands'. The report has been published in 2003, and is outdated on certain aspects. (Full text: http://www.nictiz.nl/kr_nictiz/uploaddb/download_object.asp?atoom=2128&VolgNr=1)

delivery from the individual carer to care teams dictates this change in recording habits. The often-postulated single virtual patient record that is accessible to all authorised carers is implicitly based on this ambition! The central requirement again is everyone uses and equally well understands the same language.

In the past decade we have seen a growing interest in what is called 'clinical terminology'. This term stands for the terminology needed to record all the important elements of patient care in standardised terms. This assumes that for a patient encounter significantly more will be recorded than done in current practice. In principle such a terminology is supposed to represent 'all of medicine'. If that ultimately will be one single resource covering all of medicine or a collection of compatible terminologies is a matter of growing debate. There is not an internationally accepted definition for 'clinical terminology'. It has been described as:

A clinical terminology is the collection of standard terms with their synonyms, which in the context of patient care support the recording of complaints, signs, symptoms, circumstances, process of illness, interventions, results, diagnoses, as well as the decision making of the care providers.

Descriptions in these standard terms should map unequivocally to categories in the existing (inter-) national classifications for diseases, procedures, reimbursement etc. Some argue that there should ultimately be one single global reference to support all kinds of multidisciplinary care pathways.

3 Problems with health terminology

Currently we are entering the *third phase of introducing information technology to the health care delivery system*. The *first phase* was mainly concerned with *automating the administrative process*. The *second phase* was about *delivering medical applications*. The *third phase* is about **integrating a diversity of medical and administrative systems in one coherent interoperable environment**.

At present we see **an emerging need for safe and sensible communication between these applications**. Next to a good insight of *what is really necessary to exchange*, also the *language used is important*. For instance, the slow progress in the application of knowledge-based systems may to a great deal be attributed to the lack of semantic coupling with the patient record.

Medicine is a descriptive, language intensive activity. In order to have consistent data, ideally it should be recorded only once preferably by the individual who generates it. It is therefore imperative that such data is recorded in a language that is as much as possible free of purpose. (e.g. it should not have a financial bias like many of the medical procedure classifications). The requirement from semantic interoperability for a lingua franca is that it not only enables the translation of terms, but even more that it safely communicates patient data across cultural boundaries.

The *costs of development* and perhaps more importantly the *maintenance costs of linguistic resources needed to localise clinical systems* are high. Any practical approach to the management and exploitation of linguistic resources in large-scale clinical information systems must be based on common methods and internal representations for linguistic information. This information must be reusable across a wide range of systems and local variants of those systems, and the cost of maintaining that information must be separable from those of maintaining the rest of the information system.

Significant regional differences in linguistic usage exist even within single languages, and even more so when minority languages are taken into consideration. Not only are we faced with differences in language, but also there are differences in culture. Though the Dutch and Flemish world share most of their words, the actual daily meaning of these words may be markedly different! To be truly successful, a *programme of linguistic engineering in medicine must have a strategy for recognising and managing regional as well as national linguistic and cultural differences*. These considerations further complicate and increase the *expense of 'localisation' of products for the European market*.

There is an *increasing awareness of the crucial role of a 'common medical language' for the further development of medical information systems*. The time of 'black box thinking': 'Natural Language Processing will solve the problem' is over. Moreover, NLP also needs strong and large domain models in order to be effective and reliable. Introduction of knowledge-based systems is severely hampered by the lack of common terminology. Still it must be considered that the developments are in concordance with the needs of clinical practice. The big challenge here is a technology to bring simple solutions to simple needs, but have the power to scale in an evolutionary fashion to more complex tasks at a later stage.

Traditionally classification systems have been developed for a number of quite different purposes: Statistics (ICD, etc.), Nursing (ICNP, NANDA, etc.), Human Functioning (ICF), Case Mix (DRG, etc.), Procedures/Reimbursement (ICPM, CDAM, CPT etc.), Epidemiology (ICPC, etc.), Literature retrieval (UMLS), Quality of Care, Protocols, Medical Records (Read, SNOMED, Gabrieli). Though from this list it is clear that there are standards for some elements of the Medical Record, specifically elements that characterise the patient, such as signs and symptoms, are hardly standardised.

If terminology is to be used in automated systems, terminology development can certainly not be seen in isolation:

- First of all it requires insight in the business process of the health care organisation. How does information flow through the organisation, and between organisations? So we have to address the problem of what is generally put under the umbrella of Messaging.
- Secondly for documentation we need storage structures. As stated in the introductory lines, there is a lot of activity at present in the area of structuring the Electronic Medical Record. There clearly is a relation between terminology and the names of structures in the record. There is on the one hand a number of actors in the field of record structures (HL7, OpenEHR), on the other hand a diversity of developers of terminologies. There remains a grey area, where it is not clear who should have the lead.
- Thirdly we need the terms and a delivery mechanism.

3.1 Some current bottle necks in terminology

- Specialist groups develop code lists for (national) registries; these code lists are often incommensurable with (internationally) accepted classifications;
- Classifications are frequently developed for other purposes than registration of the primary care process;
- Specific needs within certain sectors or specialties are not met by current classifications. This leads to overlapping proprietary lists and additions to existing coding systems;
- Proper methods for controlling extensions to classification systems are missing. These methods are needed for comparability, aggregation and data exchange;
- Different sectors use different standards for reporting of the same rubrics (e.g. history, diagnoses)
- Existing classifications are often inappropriate for information exchange between sectors because there is no mapping between the classifications used;
- Existing classifications are often inappropriate for registration of the care process, as they lack a sufficient level of detail;
- Existing classifications are often inappropriate for registration of the care process, as they have not been developed specifically for a given sector (ICIDH, ...)
- Proper registration of the context of care requires a suitable record structure. This structure is missing. Some mechanisms are there, but there is still too much freedom

of implementation, which may later hamper proper exchange.

- Coding of similar information does not necessarily lead to similar results. For unequivocal registration it is mandatory that similar data be coded in the same way. Support is needed to realize that.

4 Business Requirements

Terminology is a pervasive issue in healthcare, it lives everywhere, in the contents of textbooks, scientific journals, databases, messages, health cards, clinical guidelines, elements on screens of clinical information systems and so on. The profession is not always aware of the importance to have unity in language. Unity does not mean we (or our systems) all have to speak exactly the same language. What we do need is unequivocal transcription from one place to another without information loss, or if there is information loss, we need to know which information is lost.

Healthcare Information Systems are moving rapidly from management and administrative functions towards support for clinical care, quality assurance, and resource management. The market for administrative hospital systems is rapidly becoming saturated whereas the market for clinical systems has just begun to grow. The movement towards clinical systems is motivated by at least four key trends:

1. The move towards 'evidence based care'.
2. Pressure for greater cost effectiveness and more coherent planning in health care.
3. The shift from 'doctor centred care' to 'patient centred care', which requires increasing collaboration amongst different professionals and transfer of information between individuals, wards, institutions and sites.
4. The trend away from monolithic hospital wide single-vendor solutions towards heterogeneous multi-vendor solutions.

All of these changes depend on:

- Effective capture of detailed clinical information at the point of care;
- Effective communication of clinical information between users with different needs, different perspectives, and different working practices;
- Effective communication of clinical information between heterogeneous applications.

Many studies stress the fact that capture of clinical information at the point of care is essential for new technologies to be successful. At the same time it is still very awkward for care providers to review and enter substantial amounts of data during a patient encounter.

4.1 Editorial Control

4.1.1 Influence of professional bodies

It is obvious that the terminology should match the needs of the healthcare provider. If terms are clearly recognised as part of their daily professional communication, it is more likely that the terms will be used properly. They should feel a certain ownership of that terminology. Ideally the national professional societies should have a strong position in the editorial boards. For commercial offerings, which often are black boxes, such an influence does not exist. In the case of large international suppliers the potential influence of small countries is likely to be rather limited². In this context it is worth noting that there is a fear of market dominance by the large USA suppliers of terminologies and related software products.

² As an early adopter there might be opportunities for increased influence

4.1.2 Maintenance/Training

Licensing a terminology from a third party reduces the efforts of maintenance, although a central point that collects and evaluates requests for changes, and communicates these with the terminology provider, is still necessary. It will also create a dependency from this third party, and decrease flexibility. On the other hand, maintaining a proprietary terminology will be a huge task, with corresponding costs, and less possibility for international comparison and cooperation.

In traditional classifications the update frequency of procedure classifications is markedly higher than in diagnostic classifications. For clinical terminologies the same is true, typical update frequencies are in the order of twice a year. Because applications and clinical terminology are becoming ever more intertwined, version control becomes a serious issue. It is not only updating of some lists, but terms to describe record elements, protocols, screen items to name a few might be subject to change. Before introducing a clinical terminology the change management needs to be organised and be in place!

Because of the pervasive nature of a clinical terminology many need to be trained, care providers, software vendors, and other users of health data. The amount of training, and hence the costs are dependent on how complex the terminology is delivered.

4.1.3 Expertise Centre

The sheer size of a clinical terminology, and the fact that its intended use is across sectors calls for a national competence centre for both content and implementation issues. The necessity of frequent localization demands short communication lines. Even in this electronic era it is not advisable to depend solely on a supplier abroad. In order to cover sufficient depth one should think in the order of at least 3-5 persons staff for such a function in a smaller country.

4.2 Deployment

It is current practice in industry that the supplier takes initiative in setting terms and conditions under which a certain product or service is licensed. Specifically in the software industry where one supplier delivers to many customers these terms and conditions are mostly in favour of the supplier. In the situation where substantial purchases are made, the game is different. The purchaser defines the product in terms of specifications. Quite often such a purchase goes for public tender. The end result will be a more balanced agreement of rights and obligations for both the purchaser and supplier.

In case of a large national investment on terminology, be it by purchasing off-the-shelf, or by commissioning a new development, the 'tendering model' seems most appropriate. It is likely that the sheer size of the planned investment even does prescribe public tendering because of EU regulations. In case of nations joining an organisation like the Snomed SDO, these will likely not apply. In the context of tendering (perhaps also for the SDO) there are a number of requirements to consider ensuring a high quality stable product.

5. **Safety of patient data** The first and most important requirement is safety of patient data. The discussion on patient data safety is predominantly focussed on issues like confidentiality. The quality of data, and the quality of derivatives not only matters for laboratory data, but also are equally important for terminology. The product should provide sufficiently precise terms for recording. The exact meaning should be obvious from the term itself and/or the context in which it is presented. There is a second aspect of safety specific for third generation terminologies. Those are packaged in software (Terminology Server or TeS) and do have a defined functionality. Such systems can make statements about terms it contains. If e.g. a TeS drives a reminder system, its answers should be complete and correct. Though formal proof is impossible, purchasers and suppliers should specify metrics for assessment of quality.
6. **Continuity of access to patient data** For pragmatic reasons terms are mostly stored in coded form in the patient record. For future reference of those records the terms and its links to codes are necessary. If for some reason the agreement between supplier and

purchaser stops, continued use of the terminology for retrieving existing data should be secured. Also the right to convert existing data into some other terminology should be established.

7. **Multiple suppliers** In our open market terminology should be available from multiple suppliers. This applies both to the software in which it is delivered as well as to the term sets used. It must be possible to use a term set from supplier A next to a term set of supplier B within one software environment.
8. **Specification and Certification** The terminology must be explicitly specified in terms of its objectives and how it should be implemented. The supplier must also deliver the criteria for certification of the product, and rules to define certification data. The supplier may specify additional rules and tools for safe third party modifications/extensions on his terminology. As a matter of principle a trusted external party conducts the certification. A supplier will only be held responsible as far as his product has been certified.
9. **Ownership** Ownership of a terminology should only apply to the delivered/updated term collection as a whole, and the way in which it is organised. Individual words and phrases are born from the medical community, and as such are part of the public domain in the same sense that there is no owner of the English language.
10. **Extensions and updates** Healthcare language is highly dynamic, often with needs for local variants due to different healthcare delivery systems. Purchaser must secure that he has both the rights and the tools to make corrections and extensions. Purchaser should be encouraged to make corrections and extensions available to the supplier. Purchaser should retain the right to re-use extensions elsewhere.
11. **Usage** The terminology should be available for all healthcare related work and for research. This does include alternative usage in different kinds of software. Provisions must be made to avoid infringements on rights of both the supplier as well as the parties on whose behalf the purchaser enters into the agreement.
12. **Availability of resource** Both supplier and purchaser have an interest in safeguarding their investment. License agreements are likely to be made for a longer period. The time window for a purchaser is determined mainly by the time it is expected to take to change to a different supplier. In the early stages of implementing the terminology this means a shorter contract period. Later if the business process cannot be disturbed the purchaser needs a further horizon in the contract. The supplier needs to survive, and hence is interested in continued return on investments. Availability is also the purchaser's interest.

The aforementioned requirements strongly suggest aiming at some kind of open source licence agreement. Open source will most definitely not be free of cost; a complex terminology is required, and we should not expect this to appear without significant investment!

It will not be an easy task to meet all the requirements mentioned. There are tensions between the demands of multiple suppliers and safety of patient data. Certification and multiple suppliers is also a difficult issue. The model of a single market-dominating supplier may seem attractive; the situation in the software marketplace has shown that this does not necessarily guarantee quality.

5 Economic aspects

It is difficult to estimate the costs of introducing a clinical terminology. It is not just the terminology itself, but also training and software development. If the Snomed SDO is successful, the national licence fees will be low in comparison to the previous existing situation

The actual cost of implementation will depend highly on the level of ambition. Initial estimates indicate that in the first 5 years the implementation cost is at least one order of magnitude larger than the costs for licence and translation.