

Essential SNOMED: Simplifying SNOMED CT and Supporting Integration with Health Information Models

Peter MacIsaac, MB.BS, FRACGP, MPH¹, Don Walker, MB.BS², Rachel Richesson PhD, MPH³, Heather Grain FACHI⁴, Peter Elkin, MD⁵, Jon Patrick PhD⁶

¹Terminology Central, Canberra, Australia, peter@macisaacinformatics.org;

²University of Adelaide, Adelaide, South Australia, donald.walker@adelaide.edu.au;

³University of South Florida College of Medicine, Tampa, FL, USA, richesrl@epi.usf.edu;

⁴LaTrobe University, Melbourne, Victoria, Australia, h.grain@latrobe.edu.au;

⁵Mayo Clinic, Rochester, MN, USA, elkin.peter@mayo.edu

⁶University of Sydney, NSW, Australia, jonpat@cs.usyd.edu.au

ABSTRACT

SNOMED CT (SCT) has been designed and implemented in an era when health computer systems generally required terminology representations in the form of singular pre-coordinated concepts. Consequently, much of SCT content represents pre-coordinated concepts and their relationships. In this conceptual paper the role of pre- and post-coordinated terminology expressions are considered in the context of the current development direction of Electronic Health Records and the use of communications and knowledge repositories. The move from current SCT structures to an implementation form of SCT that focuses on “atomic concepts” will support post-coordination and terminology binding to information models. This core or “essential” SNOMED CT - called SNOMED Essential Terminology (S-ET) - would be smaller in terms of core concept numbers, simpler, easier to maintain and more intuitive for implementers. Our proposed implementation form of SNOMED CT would contain only “atomic concepts” with their attendant hierarchies and relationship data. These would be supported by a strict model for representing current and future pre-coordinated concepts based on the use of an existing specific post-coordination expression, grammar, or representation. The resulting concept expressions would be post-coordinated from a smaller core of atomic components. Using definitional relationships, the proposed implementation form could equate existing pre-coordinated terms with post-coordinated representations, allowing SCT to maintain links with legacy data. A strategy for testing and implementing this approach is discussed and empirical research and feasibility testing is recommended.

INTRODUCTION

SNOMED CT (SCT) is becoming the international standard clinical terminology with a new international licensing and governance process which makes it widely accessible. The adoption of SCT by multiple countries was influenced by many published studies demonstrating its comprehensive coverage [1-4] and advanced structural features. SNOMED CT has antecedents in the College of American Pathologists family of terminologies, the UK National Health Service Read Codes. As with any living language, it has absorbed content from a number of other terminologies and classifications. SCT contains concepts and terms that describe the “language of use” as well as concepts which define the “language of meaning”[5-7]. Consequently, SCT contains many pre-coordinated concepts that have varying levels of semantic complexity alongside the component or essential concepts which are themselves the building blocks of these complex clinical expressions. While there are sound historical and ongoing pragmatic reasons for this evolutionary development, the resulting mix of concept structures makes implementation within various information models complex and prone to variation. Currently, SCT is “cluttered” with pre-coordinated terms that are incompletely defined by the internal information model that exists within SCT, making transformations between existing pre-coordinated terms and post-coordinated representations difficult to achieve. This result limits opportunities for interoperability across systems, [8] which is one of the key objectives of a controlled terminology.

This conceptual paper brings to notice issues that

are well known within the health informatics community and proposes what may, at first glance, seem to be radical surgery. This proposal is in reality an extension and combination of existing features of SCT to create a more tractable solution to support both SCT development and the art and science of terminology development. This paper is not a report of a quantitative analysis of SCT structures or experimental results of the types of change proposed. These should come later, if the fundamental proposition is believed to be sound and a potential contribution to terminology development and maintenance methods.

The computational representation of data is a combination of the use of information models and terminology. We propose a variation, restructure and extension of the current SNOMED-CT terminology to support implementation in various information models.

Using a pragmatic approach the SCT would be altered in that existing pre-coordinated concepts would be identified, flagged and then defined through linkage to their atomic concepts and relationship types. The atomic or essential concepts would continue to be placed in logical and definitional hierarchies and relationship structures and subject to the use of description logic for definition, inference, and classification purposes. Existing pre-coordinated SCT concepts would retain their identifiers and be linked to the modified terminology as “pre-defined-post-coordinated concepts”, and would be logically equivalent to any post-coordination representing the same meaning. The retention of pre-coordinated concepts and the specification of their computational definitions would allow pre-coordinated terms to be used in interface applications, as pre-coordinated terms can be useful in helping data entry to be more consistent: supporting the language of use. If users have a retrieval list of pre-coordinated concepts that have post-coordinated equivalents, application developers can encourage users to use a more consistent post-coordinated form or to use entry terms that have relationships to post-coordinated expressions using fully-defined atomic concepts.

This approach to the re-organization of SCT with the formal expression of the canonical form for pre-coordination is described as “SNOMED Essential Terminology” (S-ET), or simply “Essential SNOMED”; the name coined by Dr. Walker when first describing this approach. This paper describes the case for change in SCT representation and advantages of moving to this representation, the background to the development of this approach, a representation model for pre-

coordinated concepts, and an implementation perspective. Simple examples have been selected, not to prove the feasibility of this approach, but to illustrate the principles. The need for a more technically challenging and quantitative approach to evaluation of this proposal is recognized and discussed.

TERMINOLOGIES AND INFORMATION MODELS

It is now widely accepted that health information storage is achieved through a combination of the use of controlled terminologies and standardized data models or architecture, yet the boundaries between the models used for terminology construction and health record construction are blurred. [9-12] The HL7 TermInfo project attempted to resolve this by providing guidance on how SNOMED CT could be used in HL7 version 3 messages and data structures. [12-14]

An example of this terminology model - information model interface is the question of whether concept negation should be managed within the terminology or within the data model. Should the negation be expressed as part of the terminological unit: “no history of breast cancer”, or as different components within an information model: “history of breast cancer” + “negative”? [15] The semantics can be represented in the terminology as a pre or post coordinated concept or in a combination of the data model and terminology. The machinery to support this latter approach is contained in standard information models such as the HL7 Reference Information Model (RIM). [12] HL7’s TermInfo working group has recommended that when SNOMED is being used in HL7 V3 models, negation be managed in the terminology and that the model based approach to attaching a negation indicator be deprecated. This issue points to the need for sufficient flexibility in the management of post-coordination to allow for the transformation of concept structures and modifiers between the various options. The existence of other data and information models (e.g., CDISC) – which might develop and endorse their own guidance for use of complex terminologies such as SCT - suggest that standardization of SCT terminology use in HL7 (RIM-based) applications might not guarantee interoperability with applications using other information models. [16]

While the issue of terminology and information model interaction is somewhat independent of the way that coordination of complex concepts occurs, there is a need for both pre and post-coordinated approaches to co-exist to fully support the spectrum of information representation. It is also recognized that equivalence between pre-coordinated and post-coordinated concepts has to be established to

maintain consistency in interpretation of terminology and between concept representation using different combinations of terminology and information model binding. In the current SCT infrastructure this is achieved using computation and testing the equivalence of the canonical form of the two terminological variations. This requires that all of the atomic or component concepts and pre-coordinated concepts are fully defined - not the case in practice. Having a formal definition explicitly developed for current and future pre-coordinated concepts within SNOMED would support the recognition of equivalence. [8, 17, 18]

ISSUES WITH SNOMED-CT IMPLEMENTATION

Several studies have shown that inter-rater reliability of SCT coding is poor, at least in part due to the complexity of the SCT structure and the inconsistency of existing content. [3, 19-21] This paper proposes that a simpler, more consistent representation of SCT will reduce confusion and improve the quality of SCT implementation. This would need to be tested once working subsets of the S-ET have been developed and so examine the impact on coding consistency of the interaction between the information model and the use of differently coordinated terminology.

SCT size will certainly grow as new countries adopt it, especially when it becomes the terminology to support the many uses of coded clinical data, such as public health. Trying to keep up with the need for language of use through definition of pre-coordinated concept phrases is a recipe for "combinatorial explosion" in the size of a terminology. This is bad enough in a terminology of simple structure, yet in one of SCTs complexity and richness of function, the impact is especially significant. A key technical challenge involves keeping the terminology to a manageable size and level of complexity so that it is both maintainable and supports end users' applications. A second challenge for SCT maintenance is to allow compatibility with historical versions used by legacy applications while maintaining relevance as the core terminology resource for the current and future generations of health information systems. The model proposed in this paper will support both of these objectives.

BACKGROUND TO S-ET DEVELOPMENT

In 1999 a combined pre and post-coordinated model for a medicines terminology was proposed by two of the authors (DW and PM) for Australia, based on an architecture designed earlier by DW for a proprietary drug information service. Both pre-coordinated concepts and their contained atomic components and relationship

types were accommodated. The Essential SNOMED notion, which was initially canvassed informally within the health terminology community in 2001, was further developed following a comparative technical analysis of several terminology options that were then being investigated for use in Australian General Practice [23] and which subsequently recommended use of SCT leading in time to Australia becoming an early adopter of a national SCT license. A review of candidate terminologies at that time for use in General Practice examined several options. One terminology, DOCLE, was constructed of atomic concepts, joined by operators using a Bachus Naur Form (BNF), a standard system for representation of computable expressions using syntax or rules. [24] What was notable was the extensive use of pre-coordinated terms that were constructed from atomic elements. For example "cancer@breast" was a pre-coordinated concept for "breast cancer", yet it is constructed using a post-coordination model of atomic concepts and the location operator ".@". The process of normalization of DOCLE for inclusion in a terminology service found that a number of atomic concepts needed to be created to support existing content. Considering this approach and drawing on prior experience with the development of a medicines terminology requiring a full set of atomic elements which were combined to create fully defined pre-coordinated medicines concepts, it was postulated that the SCT terminology could be significantly simplified by creating a separate data structure for the pre-coordinated concepts where these were parsed and then described in a post-coordination grammar. [25]

DESIGN OF ESSENTIAL SNOMED

Essential SNOMED would contain a complete set of "atomic concepts" from which all other concepts could be constructed by post-coordination. These atomic concepts would be carefully crafted into their hierarchies and defined by their relationships. SNOMED CT most likely contains many - if not most - of these atomic concepts. They would consist of both primitive and fully defined concepts. The large number of pre-coordinated concepts that are not in the above group should be "flagged" in the complete SNOMED CT data structure as "predefined-post-coordinated equivalent concepts", and eventually associated with their post-coordinated defining atomic concepts using a formal post coordination or compositional syntax as is already described. This group of pre-coordinated concepts would not rely on their hierarchical position or SNOMED relationships for their definition - instead they would be defined by the compositional expression (or formalism) used to construct their atomic post-coordinated concepts

(combination of existing atomic concepts), as described above. All the existing pre-coordinated concepts in SCT could remain, along with their identifier for use where the situation required this approach. Pre-coordinated concepts being added to SCT would also follow this pattern.

For example, the pre-coordinated concept “Cellulitis of the left foot with osteomyelitis of the third metatarsal without lymphangitis” can be expressed using atomic concepts and relationship types, is shown in Table 1. The concepts and relationship concepts that comprise the definition would all be considered core Essential SNOMED content.

| Operator | Disorder | Has-FindingSite | Has-Laterality |
|----------|---------------|------------------|----------------|
| | Cellulitis | Foot | Left |
| and | Osteomyelitis | third metatarsal | |
| Without | Lymphangitis | | |

Table 1 – Definitional relationships of an existing pre-coordinated SCT concept.

The current SNOMED CT terminology model specifies relationships between concepts and terms, but does not make a distinction between post-coordinated concepts expressions and pre-coordinated concepts. We propose that this distinction be made explicitly, as a tool to assist in SNOMED-CT terminology maintenance and implementation. Figure 1 describes the way that the new architectural elements could be linked with existing S-CT structures which are represented by the three elements placed at the right hand side of the figure.

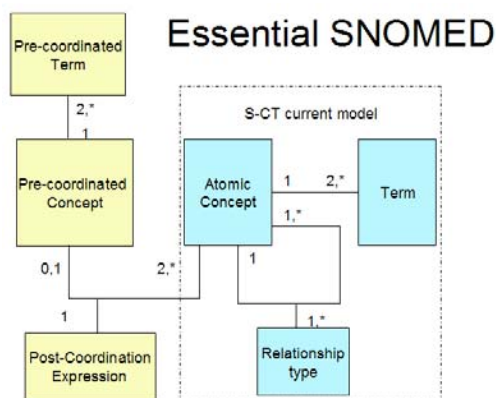


Figure 1 – Conceptual terminology model for Essential SNOMED.

DISCUSSION

At the outset it is acknowledged that this proposal is grounded in the excellent overall design and management features of SCT.

The advantages of this proposed structure for SCT are reduced size and complexity for ease of implementation and maintenance. An inevitable outcome would be a reduction in the combinatorial explosion that occurs when rampant pre-coordination of concepts and phrases occurs, yet this comes at the cost of introducing a new element in the post coordination expression that links the pre-coordinated concepts to their atomic elements. The core terminology concepts and hierarchies should be however much simplified.. The core of S-ET would grow some as new atomic concepts were added. The S-ET structure would be expressively intuitive as its approach to concept representation would support concept constructions. Hierarchical simplification would result as the definition of the many pre-coordinated-concepts would be independent of immediate hierarchies or relationships – S-ET would use the compositional expression to link with hierarchies and defining relationships of the atomic concepts. Existing approaches to canonical forms would continue and allow equivalence testing between different pre-coordinated concepts and post-coordinated expressions. Pre-coordinated concepts would still be able to be represented in a hierarchical arrangement to support inference and subsumption, however these could be calculated rather than explicit expressed as happens currently in SCT. In this model the hierarchical relationships would be inferred rather than the canonical form.

Equating pre- and post-coordination may be easier, as the computational form is actually specified for concepts within SCT. It is acknowledged that the current approach in SNOMED is not comprehensive due to incomplete set of canonical representations and possible lack of semantics to fully describe the meaning of existing semantically complex pre-coordinated concepts. Both the pre-coordinated form and the various representations of the post-coordinated concept are valid ways of describing the same concept. The first is more aligned with human interpretation and the second supports computer processing of the data. It is clear that both forms of concept representation are needed and both have to be supported by clinical terminologies such as SNOMED CT. The approach recommended in SNOMED Essential Terminology is believed to be consistent with the current SCT approach to canonical form definition.

This model is highly dependent upon an expressive and computable syntax for post-coordination. The process of moving to an S-ET distribution format will highlight any deficiencies in the current post-coordination methods and constraints as these will become explicit and subject to development. SNOMED –

CT authors can continue to develop pre-coordinated expressions if required. End users, particularly those who rely on the use of pre-coordinated concepts, will have the capacity to add locally relevant pre-coordination through a minor modification of the SCT way of managing local extensions, and in doing so would not require frequent change submissions to the core essential SNOMED terminology. As discussed earlier an S-ET model, coupled with an improved model for managing pre-coordination will support the terminology user interface.

One of the difficulties faced by SNOMED is the need to harmonize with widely used terminologies that are heavily structured on pre-coordination. LOINC and MEDCIN and most health classifications would be examples [28]. A SNOMED Essential Terminology would not need to include the pre-coordinated concepts imported from such terminologies, and could instead relate their concepts to SNOMED-CT by mapping which used the post-coordination syntax. Alternatively such pre-coordinated concepts could be placed along with existing SCT pre-coordinated content. The end result would provide the flexibility of incorporating or mapping to external terminologies, even though they may not share the same data models as SCT.

It is recognized that there are situations where pre-coordination is more efficient from a computational perspective, as in the recognition of commonly used text strings in natural language processing (NLP) applications. SNOMED Essential Terminology will allow the further development of such concepts without undue concern about the combinatorial explosion that might otherwise exist. NLP requires the consistent application of terminology and parsing of text. If a SNOMED Essential Terminology model is not adopted then it is likely that some equivalent derivative product will be created by necessity by these key application areas. Having a standard form will support consistency of output of different NLP applications.

One of the current strategies to simplify SNOMED is to restructure the relationship between terminologies and classifications. Removing or retiring classification concepts from SCT will allow them to reside in their respective classifications and have linkage to the clinical terminology by mapping or other formal constructs. SNOMED Essential Terminology proposes making a similar change to manage both the historical terminological clutter resulting from SNOMED's antecedents and use in legacy information systems. In addition it

will meet the widely accepted need to continue to manage post-coordination in a modern terminology to support the computer-human interface.

Making the transformation to S-ET

The transformation to a SNOMED Essential Terminology would require a set of suitable "relationship-types" and an appropriate post-coordination representation form or "syntax" that catered for the "pre-defined-postcoordinated equivalent concepts". SNOMED has published a BNF for this syntax. This syntax describes the core SCT concepts, and their relationships. An XML equivalent (in addition or as an alternative) may be helpful for the current computer engineering environment. This paper is not exploring the relative merits of these approaches; however the process of defining the post-coordination equivalents of existing concepts will also provide a validity check on the completeness of the syntax or post coordination model, and as such is complementary to activities of the International Health Terminology Standards Development Organisation's (IHTSDO) Concept Model Special Interest Group.

As the "pre-defined-post-coordinated concepts" could be related back to their atomic components (which are themselves part of the SNOMED hierarchy and relationship structure) it would no longer be necessary to separately define the hierarchies or associations for the pre-coordinated concepts within the terminology. This does not preclude such constructs being employed, much like current indexing activity at run-time. These hierarchies could be machine classified. For example, if the phrase "fractured ankle" was compiled from two concepts as follows:

[problem, action or issue] = "fracture"
[which has FindingSite] = "ankle"

Consequently, if it was necessary to locate "injuries of the lower limb", then the hierarchical ancestors of "fracture" would include "injuries" and those of "ankle" would include "lower limb".

The issue of what is and what isn't an atomic or pre-coordinated concept is subject to debate and the boundaries can be fuzzy. Is headache a single concept or a post-coordinated 'pain' with 'location' of 'head'? Technically, it should not be part of an S-ET based on atomic concepts but is it sufficiently common and semantically 'simple' enough to warrant inclusion? Under our proposal, "atoms refer to semantic units, not term labels or compound term labels. While it is clear that even single concepts can have compound names, it would be a conceptual error to consider a concept

such as Hodgkin's lymphoma to be a pre-coordinated concept, whereas a "fractured right femur" is patently so.

Building essential SNOMED will be necessarily a pragmatic exercise which can cope with one or other of these forms or both. The consequences of the fuzziness in determining whether an existing concept should be managed as a pre or post coordinated concepts are not expected to be significant. All are still included in SCT.

New pre-coordinated concepts could be created (if required by information systems or user preference), although this temptation may best be resisted as it is expected that the requirements for pre-coordination would become less pressing with the introduction of standard information models (e.g., HL7 V3, archetypes or OpenEHR) and the advancement of Natural Language Processing (NLP) to support data entry.[26] Complex pre-coordinated concepts can sometimes be useful in encouraging consistency in representation where small nuances may be unintentionally instantiated where no difference in clinical meaning exists. For example: Colon cancer can be represented either as: Malignant Neoplasm - Has finding site - Colon; or as: Colon - HasSpecimen - Malignant Neoplasm. This ambiguity is undesirable, and the availability of pre-coordinated concept expressions at the interface level can prevent this type of variation, and set patterns for good practice in post- coordination within terminology services.

The atomic-concepts included in S-ET would be those that are necessary and appropriate to build pre-coordinated concepts that currently exist, or may be added subsequently, as well as the atomic concepts currently in use. The boundaries around atomic concept definition are often fuzzy as discussed above. Editorial rules would be required to consider inclusion of concepts that are not "semantically atomic" but are very common. A pragmatic approach would need to be developed, and the following may suggest one strategy:

1. The entries expected to be found as defined concepts in a large medical dictionary [27]; this would likely include items that have a distinct clinical meaning and are used frequently - e.g. Lung cancer; breast cancer; direct inguinal hernia; chest pain.
2. Those concepts that cannot be adequately defined by the composition of their post-coordinated concepts due perhaps to use of an uncommon or unsupported semantic type for the relationship between elements.

As a result many of the pre-coordinated concepts found in SCT diagnoses, findings and procedures would be excluded from the atomic- concept list and be placed in the pre-coordinated group.

It is clear that any change to the structure or representation forms of SCT may have an impact on reference set (subset) development, use within value-sets, and mapping to classifications and use in local extensions. These areas need to be further examined, however S-ET would not have a significant impact, as the current SCT and S-ET would contain the same concepts and relationships. There is a significant advantage for local extensions as local terminology experts could map new local concepts to atomic elements within SCT, hence gaining the benefits of classification and relationship modelling, without having to wait for formal inclusion in later releases. The development of reference sets based on concept and hierarchy selection would also include related pre-coordinated concepts.

The feasibility of remodelling large sections of SNOMED CT, particularly when there are competing priorities for terminology development, must be assessed. While a conversion strategy has not been covered in detail, the re-organization could consider using current SCT relationships - but with some care because of their known limitations. The size of the term string, the number of individual words, the presence of relationships, and a comparison with lists of terms extracted from medical dictionaries might help identify potential pre-coordinated concepts. It is possible that a functional result to create S-ET could result from flagging pre-coordinated concepts and terms, without substantially altering the publication structure. As with most terminology development, specific tools to manage the transition to S-ET would need to be developed, refined and the end result would need appropriate checking and quality control processes and upfront attention to ongoing maintenance.

While the first efforts at instantiation of the S-ET model may involve the restructure of arbitrary twigs and branches of the SNOMED hierarchical tree, an approach proposed would be to operate on concepts identified in large sub-setting exercises where terminology of use has been identified from analysis of actual clinical use in a specific domain such as intensive care [29] or general practice .[23]

CONCLUSION

This paper has proposed a modest alteration to the structure of SNOMED CT so that it supports the co-existence of pre and post coordination in a form that advances the basic structure of what might be regarded as good terminology practice [30].

The changes do not require any fundamental changes in SCT methods, but rather a structural extension and the incorporation of existing post-coordination methods of expression into the core terminology.

This paper outlines a number of issues with the current SCT architecture and proposes a solution which is consistent with its current design and which may have a number of advantages. If the proposed model creates resonance with the end users of SNOMED CT, it should be exposed to empirical testing and considered by the IHTSDO and their related organizations. The authors hope this paper stimulates discussion and feedback. We look forward to formal testing of these ideas for feasibility and acceptance.

ACKNOWLEDGMENTS

This paper was inspired by an original paper authored by Don Walker with contributions from Peter MacIsaac arising from the "GP Vocabulary Project" supported by the Australian Department of Health and Ageing. <http://www.adelaide.edu.au/health/gp/research/current/informatics/#vocab>.

The authors would like to acknowledge the contributions of Tad McKeon, anonymous KR-MED reviewers, and numerous colleagues who have contributed in some way to this paper and our education and research on terminology and informatics.

This work has been supported in part by grants from the Centers for Disease Control and Prevention (PH00022 and HK00014; PLE).

REFERENCES

1. Wasserman, H. and J. Wang, *An applied evaluation of SNOMED CT as a clinical vocabulary for the computerized diagnosis and problem list*. AMIA Ann Symp Proc, 2003: p. 699-703.
2. Elkin, P.L., et al., *Evaluation of the Content Coverage of SNOMED CT: Ability of SNOMED Clinical Terms to Represent Clinical Problem Lists*. Mayo Clinic Proceedings, 2006. **81**(6): p. 741-748.
3. Warren, J.J. and R.P. Wilson. *Representing Cardiovascular Concepts in an Electronic Health Record Using SNOMED CT®*. in *American Medical Informatics Association Annual Symposium*. 2006. Washington, D.C.
4. Richesson, R.L., J.E. Andrews, and J.P. Krischer, *Use of SNOMED CT to Represent Clinical Research Data: A Semantic Characterization of Data Items on Case Report Forms in Vasculitis Research* Journal of the American Medical Informatics Association, 2006. **13**: p. 536-546.
5. Patrick, J. *Metonymic and Holonymic Roles and Emergent Properties in the SNOMED CT Ontology*, *Advances in Ontologies*, M.H. Orgun & Tj. Meyer (Eds). *Advances*

- in *Ontologies (AOW 2006)*, Tasmania, Conferences in Research & Practice in Information Technology, **72**, pp61-68, 2006
6. Rector, AL, Qamar R, Marley T. *Binding Ontologies and Coding Systems to Electronic Health Records and Messages*. KR-MED 2006 – Biomedical Ontology in Action. November 8, 2006, Baltimore, Maryland, USA
7. Elkin PL, Brown SH, Lincoln MJ, Hogarth M, Rector A. *A formal representation for messages containing compositional expressions*. Int J Med Inform 2003 Sep;71(2-3):89-102.
8. Andrews, J.E., et al., *Comparing Heterogeneous SNOMED CT Coding of Clinical Research Concepts by Examining Normalized Expressions*. Journal of Biomedical Informatics, 2008. **in press**.
9. Rector, A.L. *The Interface Between Information, Terminology, and Inference Models*. in *Tenth World Conference on Medical and Health Informatics: MedInfo-2001*. 2001. London.
10. Dampney, C.N.G., G. Pegler, and M. Johnson. *Harmonising Health Information Models - A Critical Analysis of Current Practice*. in *Ninth National Health Informatics Conference*. 2001. Canberra ACT, Australia.
11. Huff, S. and J. Carter. *A Characterization of Terminology Models, Clinical Templates, Message Models, and Other Kinds of Clinical Information Models*. in *AMIA Symposium*. 2000.
12. Chute, C.G., *Medical Concept Representation, in Medical Informatics. Knowledge Management and Data Mining in Biomedicine.*, Chen H., et al., Editors. 2005, Springer U.S. p. 163-182.
13. Markwell, D., *Meaning Well & Well meaning - HL7 Terminology*. 2005, The Clinical Information Consultancy Ltd. www.clininfo.co.uk
14. HL7, *Health Level Seven*. 2005, Health Level Seven, Inc.
15. Elkin, P.L., et al., *A controlled trial of automated classification of negation from clinical notes*. BMC Med Inform Decis Mak, 2005. **5**(1): p. 13.
16. Richesson, R.L. and J.P. Krischer, *Data Standards in Clinical Research: Gaps, Overlaps, Challenges and Future Directions*. Journal of the American Medical Informatics Association, 2007. **14**(6): p. 687-696.
17. Spackman, K.A. and K.E. Campbell, *Compositional concept representation using SNOMED: towards further convergence of clinical terminologies*. Proc AMIA Symp, 1998: p. 740-4.
18. Rector, A.L., et al., *The GRAIL Concept Modelling Language for Medical Terminology*. Artificial Intelligence in Medicine 1997. **9**: p. 139-171
19. Rothschild, A.S., H.P. Lehmann, and G. Hripcsak, *Inter-rater Agreement in Physician-coded Problem List*, in *American Medical Informatics Association*. 2005: Washington,

- D.C.
20. Burkhart, L., et al., *Mapping parish nurse documentation into the nursing interventions classification: a research method*. Comput Inform Nurs, 2005. **23**(4): p. 220-9.
 21. Chiang, M.F., et al. *Reliability of SNOMED-CT Coding by Three Physicians using Two Terminology Browsers*. in *American Medical Informatics Association Annual Symposium*. 2006. Washington, D.C.
 22. Andrews, J.E., R.L. Richesson, and J.P. Krischer, *Variation of SNOMED CT Coding of Clinical Research Concepts among Coding Experts*. JAMIA, 2007. **14**: p. 497-506.
 23. Walker, D. and e. al, *General Practice Vocabulary Project*. 2001, Commonwealth Department of Health and Ageing.
 24. *Welcome to DOCLE Systems*. 2008, . <http://www.docle.com.au/>
 25. Walker, D. and e. al., *Essential SNOMED, unpublished project report*, University of Adelaide.
 26. Elkin, P.E., et al., *An Evaluation of the Content Coverage of SNOMED CT for Clinical Problem Lists*. Mayo Clin Proc, 2006. **81**(6): p. 741-8.
 27. Elkin, P.L., et al., *Guideline and Quality Indicators for Development, Purchase and Use of Controlled Health Vocabularies*. International Journal of Medical Informatics, 2002. **68**(1-3): p. 175-186.
 28. Rosenbloom ST et al. *Using SNOMED CT to Represent Two Interface Terminologies* In review
 29. Patrick J, Herkes R, Ryan A. *Enhancement technologies for clinical information systems*. HISA NSW conference proceedings 2008.
 30. Cimino JJ. *Desiderata for controlled medical vocabularies in the twenty-first century*. Methods Inf Med 1998;37(4-5):394-403