# Ontological Issues in using a Description Logic to Represent Medical Concepts: Experience from GALEN: Part 1 - Principles

**Alan Rector and Jeremy Rogers**

**Medical Informatics Group**
**Department of Computer Science**
**University of Manchester**
**Manchester M13 9PL**
**Tel +44-161-275-6188/6239/7183**
**FAX: +44-161-275-6204**
**email <rector/jeremy>@cs.man.ac.uk**

**& *Open*GALEN**
**www.opengalen.org**

**Abstract**

Concept models, or 'ontologies' represented in description logics are becoming central to many efforts in re-usable medical terminologies. Developers of any ontology must make a series of choices concerning how concepts are represented. These choices are made more difficult by the limitations of the logical formalisms in which ontologies are represented and the complexity, size and scope of medicine. Principles for ontology structure to guide developers are still poorly understood. This paper articulates the principles arrived at from the experience of the GALEN Programme. The complete GALEN Common Reference Model is available from *Open*GALEN at www.OpenGALEN.org

Keywords: Terminology, Ontology, Knowledge Representation, conceptual modelling

## 1. Introduction: Goals and criteria for the success of the ontology

GALEN seeks to provide re-usable terminology resources for clinical systems. The heart of GALEN is the use of an ontology, the Common Reference Model, formulated in a specialised description logic, GRAIL [1]. There are several other efforts to use description logics or closely related formalisms in medical terminology including SNOMED-RT [2-4], the pioneering work of the 'Canon group' [5-7], and more language oriented work, *e.g.* [8, 9], plus less formal methods of representing the semantics of terminologies used in the NHS Clinical Terms (Read Codes) [10]. Ontologies based on related formalisms play a key role in attempts to develop re-usable problem solving methods by Musen and his colleagues in [11-13]. Outside of medicine, description logics are increasingly being used for knowledge representation, indexing, and data management in a wide variety of fields, *e.g.* [14-18] and for encoding 'meta knowledge' for the Web. [19, 20]

This paper presents a unified approach to the GALEN ontological schemas, previously partly described in [21, 22] which is just one part of the overall GALEN architecture described in [23-25] High level issues related to Cimino's desiderata for terminologies [26] can be found in [24]. Further information on the ontology described and all these topics can be found on the *Open*GALEN web site, www.opengalen.org and the subsequent paper [27].

The overall aim of the *Open*GALEN terminology resources is to support clinical information systems – to allow information to be recorded in electronic patient records, abstracted from them, re-organised to provide clearer views of an individual patient, aggregated for management, research, and administration, and linked to knowledge resources – decision support, bibliographic, and general web-based information systems. Classification is the key to abstraction, re-organisation, and re-use. The key technical criterion for the GALEN ontology is therefore: *whether its definitions and descriptions are correctly and completely classified.*

More generally, we can describe any ontology in terms of:

a) *Expressiveness* – the ability to represent formally the concepts clinicians use: symptoms, diseases, procedures, etc.

b) *Classification* – the ability to infer the correct classification (indexing) of the concepts represented.

GALEN's idealised goal was an ontology which could express 'all and only' what is medically sensible. In practice, the 'all' has taken precedence over the 'only', and the goal has been to have an ontology which is sufficiently expressive to capture the statements and abstractions used by clinicians and classify them correctly while rejecting patent nonsense.

This is essentially a technical, logical and software oriented goal. The GALEN ontology is designed as an internal 'assembly language' rarely to be seen directly by users or even most software developers. Intuitive user-oriented presentation is handled separately [23, 28].

## 2. Basic Principles

### *2.1 'Logical approximations', labels, & not 'lying to the system'*

Any logical model for knowledge representation is at best an approximation of concepts as used in human language and thought. A 'logical approximation' may seem an oxymoron, but logical models of any kind behave very differently from language or 'units of thought'. The fluid dependence on context typical of thought and language eludes logical representations. More specifically: :

a) There are well known trade-offs between expressiveness and computational tractability in logical systems [29, 30].

b) Reality is fractal – no matter how much detail is represented in the model, it is always possible to represent more.

Since any ontology is an approximation, the labels attached to concept representations internally in the ontology are at best also approximations. Arguments such as "Is the hand still a division of the upper extremity if its been amputated?" or "Is there a difference between an 'act' and a 'deed'?" are largely futile. When arguments over the labelling of concept representations occur, GALEN asks two questions:

a) Does the representation represent *some* concept which most agree is useful and clearly defined even if they cannot agree on what to call it?

b) Is the label seriously misleading? Ambiguous? Does it mean different things to different groups?

On the whole GALEN has found that non-understanding is better than misunderstanding. Hence internal labels often appear awkward, *e.g. IntrinsicallyPathologicalPhenomenon*.

However, the restrictions required to make logical approximations computationally tractable raises a more serious problem than mere labelling. Logic guarantees only that true conclusions follow from true premises. If the logical formalism is not expressive enough, then users may ‘work around it’ to make statements which are not just approximate but false – they may ‘lie to the system’ [30]. Even if the result serves well in one application its behaviour for other applications will be difficult to predict. The most common example of ‘lying to the system’ is the use of ‘is kind of’ instead of ‘is part of’’ which works for disorders and procedures but breaks down for more general descriptions of pathophysiology (see [24]).

### *2.2 Categories, instances, and natural kinds*

The GALEN concept model contains only *categories* (also known as ‘types’, or ‘classes’). Statements about the real world in medical records represent *instances* (‘individuals’) of these categories. Categories can be abstract, such as ‘Phenomenon’ or ‘Disease’, general such as ‘Diabetes’ or very specific such as ‘aspirin’ or ‘foot’. Categories can be specialised – *e.g.* “aspirin” can be specialised to “children’s aspirin” “foot” can be specialised to “left foot”, “deformed foot”, “deformed left foot”, etc. Instances can only be described but not specialised. It makes no sense to say “a kind of this tablet of Aspirin” or “a kind of Alan’s foot”.

Many ontologies identify instances as being concepts specialised to the level of detail required for a particular application (see Brachman’s ‘Living with Classic’ [31]). This approach is fatal to re-use, since, as Brachman so elegantly demonstrates, the appropriate level of detail for one application will almost certainly be different from that for another. However, GALEN must still determine which categories should be represented as ‘elementary’ and which should by defined *i.e.* be ‘composite’. There are two issues in deciding this issue for any given concept:

a) Whether it is *possible* to define the concept. A definition must give necessary *and sufficient* criteria for recognising that concept. Many important concepts defy definition by sufficient criteria. Such concepts are termed ‘natural kinds’ and include most simple notions such as “leg”, or “tree”, “process”, “flow”, etc. Natural kinds can also occur at a more abstract level. One might be tempted to define “heart valve” equivalently to “valve in the heart”, and “valve” as a “structure which functions as a valve”. But this combination results in the “foramen ovale” being classified as a kind of “heart valve”, since it is undoubtedly located in the heart and functions as a valve. Hence, in GALEN all named body parts are natural kinds.

b) Whether it is *useful* to define the concept with respect to the needs of the applications to be supported by the model. Some concepts are simply not worth the trouble to define even if definition might be possible.

### *2.3 Explicitness and orthogonal taxonomies*

Potentially, it should be possible re-arrange the ontology along any axis. In a description logic this corresponds to saying that it should be possible to classify any concept representation on any of its properties independently. Therefore, all properties must be represented explicitly and independently, even at the cost of apparent redundancy. For example, the indications for a drug must be represented separately from its actions even though one can often be inferred from the other, *e.g.* that an indication of 'relief of brochoconstriction' be represented separately from an action of 'bronchodilatation'.

GALEN formulates this principle as the 'principle of orthogonal taxonomies' [32, 33]. In summary, there are two criteria:

a) Every elementary concept must have one, and only one, elementary parent. All multiple classification must be inferred from descriptions and definitions.

b) The elementary children of an elementary concept are considered disjoint but non-exhaustive – *e.g* if the children of 'Bone' are a list of specific bones and categories of bones, the bones and subcategories of bones must not overlap, but it cannot be assumed that the list is complete.

Our experience is that the principle of orthogonal taxonomies is key to re-use.

### *2.4 Normalisation and canonical forms*

Normalisation is the process of reducing an expression to a standard, usually simplest, form. There are two levels of normalisation which must be dealt with:

a) Logical – *e.g.* to reduce "long bone fracture located in the femur" to "fracture of femur". This is a purely logical operation dependent on the representation of "fracture", "bone", "long bone", and "femur".

b) Pragmatic – *e.g.* to agree whether a particular procedure is to be represented as the equivalent of "fixation of femur by means of insertion of pins" or "insertions of pins to fixate femur" [34]. In this case there is no simple logical solution and the problem must be solved by conventions and guidelines which are outside the formal model.

## 3. Technical issues for the design of an ontology

### 3.1 First class entities and modifiers

The concepts in ontologies tend to divide into two kinds:

a) Those which represent things which can exist on their own, *e.g.* physical objects, processes, ideas, etc.

b) Those which only make sense when linked to some other object as a modifier, modalities, or collections. Modifiers are notions such as "severe", or "soft", or "short" which describe first class concepts and specialise them further. Modalities are notions such as presence, uncertainty, family history etc. which take their meaning from the kernel first class concept. Sowa [35] after Pierce terms the first class objects simply 'firsts' and the modifiers 'seconds'.

The most important practical difference between first class entities and modifiers in a clinical ontology is that:

a) Lists of first class entities are almost always 'open', *i.e.* they cannot be assumed to be complete so that it is not legitimate to infer from the negation of some that one of the others is present. Lists of diseases, operations, and even bones and joints, are almost never complete. Reality is fractal and more detail can almost always be added.

b) Lists of modifiers may be 'closed', *i.e.* may be assumed to be complete so that inference of the form "not raised or normal, therefore depressed" can be justified logically, though they must be used with care clinically.

GALEN treats all concepts as 'open'. It never makes inferences such as "not absent implies present" on the grounds that they risk imputing a degree of logical rigour to clinician's statements which is rarely intended.

### 3.2 Nominalised relations or 'Features'

The choice of what should be a semantic link and what should be a concept is less simple than it seems, since any semantic link can be nominalised into a concept, *e.g.* in GRAIL notation:

*Disease <u>which</u> hasSeverity severe*

the link can be nominalised to give an expression of the form:

*Disease <u>which</u> hasFeature (Severity <u>which</u> hasState severe)*

Given that this transformation is always possible, at one extreme a system could be built with just two semantic links – hasFeature and hasState – so how should the decision be made? There are two criteria:

a) Need for further description. In most formalisms, semantic links can be organised into a kind-of hierarchy, but they cannot themselves be described. Therefore, if the 'fact of being linked' may need to be described, even if in only some cases, then the link must be nominalised to a feature.

b) Consistency of representation. If there are a series of properties which appear analogous, it is almost impossible to maintain a system in which some are represented as links and some as features, therefore if any must be described as in a), all should be nominalised.

### *3.3 Dualities and qua-induction*

If we define "mother" as the "female parent of a child", we imply the possibility of describing a "child with a female parent". (This process is sometimes called 'qua induction'.) Many medical concepts come naturally in such dualities, and it is not always obvious which should be represented as primary. For example the 'process of ulceration' produces 'ulcer lesions'. Should the process be defined in terms of the lesion or vice versa? or should both be treated as elementary and related? or should both be treated as aspects of some broader abstraction which we might phrase "ulceration situation"? The choice is unclear, but it needs to be made consistently. GALEN has chosen to have just two concepts in all such cases corresponding to the process and resultant lesion, to make the process elementary, and to define the lesion in terms of the process.

### *3.4 Top level ontologies*

The original belief of those developing the GALEN ontology was that it would be built from the bottom up. The top level ontology was seen as making little difference to classification and inference. Experience has largely confirmed this view technically but, paradoxically, refuted it pragmatically with respect to the development process. Regardless of its effect on inference, an agreed top level schema has proved essential to allow groups to co-operate effectively.

However, just as all ontologies are approximations, all high level ontologies are to some degree arbitrary. Our conceptualisation of the world does not break down into a sequence of disjoint partitions. There are several existing starting points – PENMAN [36], Cyc [17, 37], traditional schemes deriving from Shank [38] and others in the AI/Linguistics field, and those deriving from Pierce via Sowa [35]. GALEN's was originally adapted from that of Cyc [17].

In addition, each major field such as medicine requires one or two very high level abstractions which are very broad disjunctions which cut across the traditional boundaries of top level ontology. In GALEN, the concept representation *Phenomenon* and attribute *involves* are designed to cover anything which is or might become a disease, disorder, or condition.

## 4. Ontological issues which interact with the semantics of the formalism

### *4.1 Negation and uncertainty*

Negation and uncertainty pose difficulties for at least four reasons:

a) The meaning of negation and uncertainty in clinical medicine is unclear. For example, if there is no mention of diabetes in the record at all, what should be the answer to a query "Does the patient have diabetes?" Most database systems would answer "no" on the basis of a 'closed world assumption' and 'negation as failure' – the assumption that all relevant information about the domain of discourse is contained in the database and that therefore failure to find a fact can be taken as equivalent to its negation. In many clinical applications, neither assumption seems safe. Furthermore, if uncertainty is catered for, should it be included with negation or be a separate dimension? *e.g.* what are the comparative meanings of "possibly present" and "possibly absent"? .

b) The scope of the negation is often unclear. At least three cases must be distinguished: a) "It is not the case that the patient has X"; b) "The patient has non-X", and c) "The patient has X but not some specific kind of X". Cases a) and c) are the common ones in medical records.

c) Adding negation and uncertainty to formalisms increases their computational complexity and makes normalisation to a unique canonical form difficult. (Note that even if the formalism supports negation, the designers of the ontology may decide not to use it. KRSS which underlies SNOMED-RT supports negation, but it is the authors' understanding that negation is not used in the SNOMED ontology.)

d) Negation and uncertainty are often represented in the information systems which hold the concept representation from the ontology. If negation can be represented both in the information system and the ontology, then the meaning of all possible combinations of negation in the two systems must be defined.

GALEN uses a formalism without negation and includes constructs such as ‘presence’ and ‘absence’ which can be qualified by a certainty.

### *4.2 Defaults and indexing*

The definition for ‘B is a kind of A’ in all formal logical representations is that ‘All Bs are kinds of A’. Hence all of the properties in the definition and descriptive axioms about ‘As’ must apply to all ‘Bs’ without exception. Adding exceptions to such logical schemas has had little success’ [39]. Therefore, in GALEN or any other system based on description logic, statements in definitions and descriptions are ‘indefeasible’– *i.e.* must apply to all subconcepts without exception. This contrasts with the frame systems from which such representations were derived, in which default values for a slot may be over-ridden.

However, if additional facts are indexed by a description-logic-based ontology, that ontology can still help with reasoning about defaults by gathering together the set of all the most specific candidate values indexed. (One candidate value is more specific than another if the concept from which the more specific is inherited is a kind of the concept from which the less specific value is inherited [40]). For example, a drug ontology can be used to index potential side effects, even though some side effects are subject to exceptions [41]. GALEN refers to such indexed statements as ‘extrinsic’ because they do not affect classification and therefore are not part of the ontology proper.

GALEN’s experience is that if the taxonomies are properly orthogonal, the set of candidate values will usually have exactly one member. If it has more than one member, then GALEN treats this as a signal that other reasoning methods and knowledge are required.

### *4.3 Normative statements, congenital malformations, and imputed intentions*

Many of the descriptive axioms used in terminology models are actually ‘normative’ rather than absolute, *i.e.* they really pertain to our view of ‘normal’ anatomy, physiology etc. This gives problems when describing congenital malformations and mutilations. There are two complementary approaches to this problem:

a) To adjust the interpretation of the semantic links and concepts. For example, GALEN interprets the hasDivision link in such a way that ‘Hand isDivisionOf Arm’ is true, even if the hand is severed from the arm. Since we can still talk about the missing hand, this is the best ‘logical approximation’.

b) To model both normal and abnormal but use the interface to limit the normal view only to the normal conformation.

Normative statements give more difficulty when applied to procedures and treatments. To take the classic example used by O'Neil [34] "Insertion of pins in the Femur". This operation is almost certainly performed in order to fixate the femur. To classify this operation under "Operations to fixate long bone" this information must be added to the description of this category of operations. However, to do so risks imputing unstated intentions to the clinicians. GALEN is cautious about making normative statements in the Common Reference Model, but finds some cannot be avoided.

### *4.4 The attribute hierarchy*

In GRAIL, as in many but not all description logics, one attribute (role) can be a kind of another, just as one category can be a kind of another. There are several key uses for the attribute hierarchy, but three deserve special mention and are supported by naming conventions in GALEN:

a) To allow single-valued and multi-valued variants of an attribute. Logically, the single-valued attribute must always be a descendent of the multi-valued attribute, and its presence is signalled by the naming convention that the single-valued variant contains 'specific' or 'specifically', e.g. hasSpecificConsequence, or actsSpecificallyOn. These 'specific' attributes are often used to indicate a main, or primary action, cause, etc. – see "multiple causation".

b) To allow direct and transitive variants of an attribute. In this case facts almost always involve the 'direct' variant and queries the transitive variant.

c) To allow very general queries, such as "disorders of heart". The *involves* attribute (see "constructive attributes" below) is the best example of this.

### *4.5 Transitive relations*

Parts and wholes, causal links, and connections are all transitive and some characteristics are 'inherited' along these transitive relations. Establishing the pattern of transitive relations and inheritance along them is a key part of any ontology of medicine [1]. GALEN also uses transitive relationships for a number of other purposes, including:

a) In the representation of syndromes, so that the presence of a syndrome is a kind of the presence of each disease in the syndrome.

b) In the representation of procedures, so that a global procedure acts on all of the structures acted on by its subprocedures.

### *4.6 Concept constancy in compositions and avoiding 'Up-to-ism'*

If a concept's representation depends on its use, then this limits re-use. Concepts such as "lobe of the liver" or "fluid in cyst in the kidney" should appear the same regardless of context – whether as aspects of disease, targets of surgery, substance to be injected or drained, or specimen in a pathology examination. Since many of these concepts are themselves composite, GALEN must allow definitions to be embedded within definitions to any degree required – *i.e.* GALEN supports expressions such as "upper part *of* third segment *of* middle lobe *of* right lung" to any depth and their subparts appear the same in the composition as when isolated.

Arbitrary limits on the complexity of expressions also limit re-usability. Arbitrary limits are sometimes caricatured as 'up-to-ism' – you can have 'up to' some limit of a particular option. Traditional enumerated medical terminologies have suffered seriously from 'up-to-ism' with limits on the depth of the hierarchy and the number of items at any given level. The most common source of 'up-to-ism' in logical knowledge representations in medicine arises from the handling of locations and sites, *e.g.* Location-right lung, sublocation-middle lobe, sub-sub location-third segment. GRAIL's support for transitive relations means that definitions avoid the need for such constructions.

### *4.7 Adjacency, wholeness, and spatial (and temporal) reasoning*

"A fracture of the tibia and fibula" makes sense; "a fracture of the tibia and humerus" does not. Furthermore, "a fracture of the tibia" is not a kind of "fracture of the tibia-and-fibula-combined-structure" – which it would be by GALEN's rules if we simply considered it as a single structure with parts "tibia" and "fibula". In an analogous case, "an inflammation of the ascending colon" is a kind of "inflammation of the large bowel" but not of an "inflammation of the large bowel as a whole". Such reasoning is beyond that available in current description logics although there are several experimental 'work arounds'.

### *4.8 Numerical conversions, calculations, and other 'non-terminological' reasoning*

There are numerous services that a user might naturally expect to be packaged with a terminology but which require entirely different reasoning from logical classification based on definitions and descriptions. The most obvious of these is conversion between different unit and co-ordinate systems. These require additional reasoners packaged within a 'terminology server'.

## 5. Discussion

In establishing the principles outlined in this paper, GALEN has been guided by earlier work, but has attempted to make choices based on evidence. In this context evidence may be of two kinds:

a) That the classifications inferred were correct with respect to the applications expected

b) That the ontology itself could be understood and maintained

Distinctions were motivated by technical considerations and the limitations of logical representations. Presentation to users was delegated to 'intermediate representations' and 'perspectives' [23]. The structure of part-whole relationships, and transitive relations more generally, has required a series of distinctions between divisions, layers, components, etc. which are not obvious. Likewise, the decision about which information should be represented in the ontology itself, and which should be indexed by the ontology through 'extrinsic' links, depends on whether or not a given construct can be treated as indefeasible for all contemplated applications or whether its values must be treated as a default subject to exceptions. This is often as much a question of software engineering as knowledge representation.

Fundamental to the approach is the 'principle of orthogonal taxonomies' which forces developers to make as much information as possible explicit. This principle has repeatedly shown its value, most recently in the development of the drug ontology for the UK PRODIGY project [42]. Other key internal criteria of the approach are 'object constancy', 'not lying to the system', avoiding 'up-to-ism', and a conservative approach to handling negation. Making it work at all requires accepting the limitations intrinsic to the fundamental fact that any logical representation is, at best, an approximation of our conceptual and linguistic systems.

These principles are far from a complete solution to all the problems encountered. At least seven major issues remain outstanding:

a) Improved pragmatic normalisation where logical criteria do not suffice.

b) Improved handling of congenital abnormalities

c) Improved handling of 'normative' statements

d) Improved handling of other patterns for the interaction of transitive semantic links

e) Improved handling of adjacency and other spatial and temporal notions

f) How to best take advantage of the improvements in description logic technology now becoming available [43, 44].

Issues a) through d) concern primarily the ontology itself, although ‘pragmatic normalisation’ (a) may be as much a matter of improved tools as improved representations. Issues e) and f) are more closely related to the underlying description logic formalism and how to take best advantage of the extensive advances in the technology since GALEN was conceived.

The subsequent paper [27] examines the GALEN schemata for the top level ontology, anatomy and diseases and relates the concrete problems encountered to the principles articulated in this paper. Separate papers describe the schemata for surgical procedures [28] and drugs [42].

## Acknowledgements

The research reported was supported in part by the European Healthcare Telematics programme.  The partners in the GALEN and GALEN-IN-USE projects have contributed extensively.  Particular mention is due to Anthony Nowlan who first formulated key parts of what has become the GALEN ontology.

## References


1. Rector A, Bechhofer S, Goble C, Horrocks I, Nowlan W, Solomon W. The GRAIL concept modelling language for medical terminology. Artificial Intelligence in Medicine 1997;9:139-171.
2. Spackman KA, Campbell KE, Côté RA. SNOMED-RT: A reference Terminology for Health Care. Journal of the American Medical Informatics Association (JAMIA) 1997((Symposium special issue)):640-644.
3. Spackman KA, Campbell KE. Compositional concept representation using SNOMED: Towards further convergence of clinical terminologies. Journal of the American Medical Informatics Association 1998(1998 Fall Symposium special issue):740-744.
4. Campbell KE, Tuttle MS, Spackman KA. A "lexically-suggested logical closure" metric for medical terminology maturity. Journal of the American Medical Informatics Association 1998(Fall Symposium Special Issue):785-789.
5. Evans DA, Cimino J, Hersh WR, Huff SM, Bell DS, The Canon Group. Position statement: Towards a medical concept representation language. Journal of the American Medical Informatics Association 1994;1(3):207-217.
6. Cimino J. Controlled Medical Vocabulary Construction: Methods from the Canon Group. Journal of the American Medical Informatics Association 1994;1(3):296-197.
7. Tuttle MS. The position of the canon group: A reality check. Journal of the American Medical Informatics Association 1994;1(3):298-299.
8. Hahn U, Romacker M, Schulz S. How knowledge drives understanding - matching medical ontologies with the needs of medical language processing. Artificial Intelligence in Medicine 1999;15(1):25-52.
9. Zweigenbaum P, Bachimont B, Bouaud J, Charlet J. Issues in the structuring and acquisition of an ontology for medical language understanding. Methods of Information in Medicine 1995;34(1/2):15-24.
10. Brown P, O'Neil M, Price C. Semantic definition of disorders in Version 3 of the Read Codes. Methods of Information in Medicine 1998;37:415-419.
11. Musen M. Modern architectures for intelligent systems: reusable ontologies and problem-solving methods. Journal of the American Medical Informatics Association 1998(Symposium supplement):46-54.

12. Erikson H, Shahar Y, Tu S, Puerta A. Task modelling with reusable problem-solving methods. Artificial Intelligence 1995;79:293-326.

13. Tu S, Eriksson H, Gennari J, Shahar Y, Musen M. Ontology-based configuration of problem-solving methods and generation of knowledge-acquisition tools - application of PROTEGE-II to protocol-based decision-support. Artificial Intelligence in Medicine 1995;7:257-289.

14. Goble C, Haw C, Bechhofer S. Describing and classifying multimedia using the description logic GRAIL. In: Proceedings of the SPIE Conference on Storage and Retrieval of Still Image and Vidio IV, SPIE vol 2670; 1995; San Jose; 1995. p. 132-143.

15. Guarino N. Understanding, building and using ontologies. International Journal of Human Computer Studies 1997;46:293-310.

16. Borgida A. Description logics in data management. IEEE Transactions on Knowledge and Data Engineering 1995;7(5):671-682.

17. Lenat DB, Guha RV. Building Large Knowledge-Based Systems: Representation and inference in the Cyc Project. Reading, MA: Addison-Wesley; 1989.

18. Fikes R, Farquhar A, Pratt W. Information brokers: Gathering information from heterogeneous information sources. In: Stewman JH, editor. Ninth Florida Artificial Intelligence Research Symposium (FLAIRS-96); 1996; Key West, Florida; 1996. p. 192-197.

19. Motta E, Buckingham-Shum S, Domingue J. Ontology-Driven Document Enrichment: Principles and Case Studies. In: Knowledge Acquisition Workshop; 1999; Banf, Canada; 1999. p. http://sern.ucalgary.ca/ksi/kaw/KAW99/papers.html.

20. Fensel D, Decker S, Erdmann M, Studer R. Ontobroker: How to enable intelligent Access to the WWW. In: Integration AWoK, editor.; 1999; 1999. p. www.aifb.uni-karlsruhe.de/WBS/dfe/iii99.html.

21. Hardiker NR, Rector AL. Modeling nursing terminology using the GRAIL representation language. Journal of the American Medical Informatics Association 1998;5:120-128.

22. Rogers J, Rector A. The GALEN ontology. In: Brender J, Christensen J, Scherrer J-R, McNair P, editors. Medical Informatics Europe (MIE 96); 1996; Copenhagen: IOS Press; 1996. p. 174-178.

23. Rector AL, Zanstra PE, Solomon WD, Rogers JE, Baud R, Ceusters W, et al. Reconciling Users' Needs and Formal Requirements: Issues in developing a Re-Usable Ontology for Medicine. IEEE Transactions on Information Technology in BioMedicine 1999;2(4):229-242.

24. Rector AL. Clinical Terminology: Why is it so hard? Methods of Information in Medicine 1999;38:(in press).

25. Rector A, Solomon W, Nowlan W, Rush T. A Terminology Server for Medical Language and Medical Information Systems. Methods of Information in Medicine 1995;34:147-157.

26. Cimino J. Desiderata for controlled medical vocabularies in the twenty-first century. Methods of Information in Medicine 1998;37(4-5):394-403.

27. Rector AL, Rogers JE, Solomon WD. Ontological issues in using a description logic to represent medical concepts: Part II: The GALEN high level schemata. Methods of Information in Medicine 2000:(in press).

28. Rogers J, Solomon D, Rector A, Zanstra P. From rubrics to dissections to GRAIL to classifications. In: Medical Informatics Europe (MIE-97); 1997; Thesalonika, Greece: IOS Press; 1997. p. 241-245.

29. Brachman R, Levesque H. The tractability of subsumption in frame-based description languages. In: AAAI-84; 1984: Morgan Kaufman; 1984. p. 34-37.

30. Doyle J, Patil R. Two theses of knowledge representation: Language restrictions, taxonomic classification and the utility of representation services. Artificial Intelligence 1991;48:261-297.

31. Brachman RJ, McGuinness DL, Patel-Schneider PF, Resnick LA, Borgida A. Living with Classic: When and how to use a KL-ONE-like language. In: Sowa J, editor. Principles of Semantic Networks: Explorations in the representation of knowledge. San Mateo, CA: Morgan Kaufmann; 1991. p. 401-456.

32. Rector A. Coordinating taxonomies: Key to re-usable concept representations. In: Barahona P, Stefanelli M, Wyatt J, editors. Fifth conference on Artificial Intelligence in Medicine Europe (AIME '95); 1995; Pavia, Italy: Springer; 1995. p. 17-28.

33. Rector A. Thesauri and formal classifications: Terminologies for people and machines. Methods of Information in Medicine 1998;37(4-5):501-509.

34. O'Neil M, Payne C, Read J. Read Codes Version 3: A user led terminology. Methods of Information in Medicine 1995;34:187-192.

35. Sowa J. Knowledge Representation: Morgan Kaufmann; 1999.

36. Bateman JA. Upper modelling: a general organization of knowledge for natural language processing. Penman Development Note: USC/Information Sciences Institute; 1989.

37. Guha R, Lenat D. Enabling agents to work together. Communications of the ACM 1994;37:127-142.

38. Schank RC, Abelson R, P. Scripts, Plans, Goals, and Understanding. Hilsdale, NJ: Lawrence Erlbaum Associates; 1977.

39. Etherington D. Formalising nonmonotonic reasoning systems. Artificial Intelligence 1987;31:41-85.

40. Touretzky D. The Mathematics of Inheritance Systems. Los Altos, CA: Morgan Kaufmann; 1986.

41. Solomon W, Heathfield H. Conceptual modelling used to represent drug interactions. In: Barahona P, Veloso M, Bryant J, editors. Twelfth International Congress of the European Federation for Medical Informatics, MIE-94; 1994; Lisbon, Portugal; 1994. p. 186-190.

42. Solomon DS, Wroe C, Rogers JE, Rector A. A reference terminology for drugs. Journal of the American Medical Informatics Association 1999(1999 Special Conference Issue).

43. Horrocks I. The CAMELOT Web Page. In:; 1998.

44. Horrocks I. Using an expressive description logic: FaCT or Fiction. In: Cohn AG, Schubert LK, Shapiro SC, editors. Principles of Knowledge Representation and Reasoning: Proceedings of the Sixth International Conference on Knowledge Representation (KR 98); 1998; San Francisco, CA: Morgan Kaufmann; 1998. p. (in press).