Viewpoint ■

The Unified Medical Language System:
Toward a Collaborative Approach for Solving Terminologic Problems

KEITH E. CAMPBELL, DIANE E. OLIVER, EDWARD H. SHORTLIFFE

A b s t r a c t   The approach taken by the Unified Medical Language System (UMLS), in which disparate terminology systems are integrated, has allowed construction of an electronic thesaurus (the Metathesaurus) that avoids imposing any restrictions upon the content, structure, or semantics of the source terminologies. As such, the UMLS has served as a unifying paradigm by providing appropriate links among equivalent entities that are used in different contexts or for different purposes. It accordingly provides a vehicle through which possibly orthogonal semantic models can co-exist within a single framework. This framework provides a model for the collaborative evolution of biomedical terminology and allows a synergistic relationship between the UMLS and its source terminology systems.

■ **JAMIA.** 1998;5:12–16.

Informatics professionals have diverse perspectives and approaches for solving contemporary informatics problems. Maintaining such diversity can be a productive strategy for overcoming obstacles when an optimal method has yet to be demonstrated. To provide general benefit to those who draw on the work of the informatics community, however, developmental efforts need to be reconciled periodically by identifying where approaches are similar, and how and why they diverge. By recognizing the similarities, and by working cooperatively on common tasks while independently pursuing truly unique approaches, researchers benefit by their ability to leverage one another's work. At the same time, consumers benefit by the availability of integrated applications that take advantage of the strengths of the diverse approaches.

The 10-year anniversary of the National Library of Medicine's Unified Medical Language System (UMLS) represents an opportunity for such reflection about diverse perspectives and approaches used in constructing terminologies for biomedicine. We believe that the UMLS provides a framework within which researchers and developers can reconcile common aspects of their approaches, while ensuring that there is an opportunity within this framework to explore unique approaches and to meet specialized needs.

Collaboration: The Real Grand Challenge

Many informatics researchers recognize that "solving" the terminology problem is viewed by many observers as a litmus test of the credibility and role of the medical informatics community. Sittig[1] surveyed informatics researchers and found that an effective approach to the terminology problem was identified as one of the grand challenges for the next decade. There is little doubt, however, that if the same question had been asked 10 years earlier, the development of a standardized biomedical terminology would have been identified as a key goal at that time as well.* How can

*The initiation of the UMLS project 10 years ago is evidence to support this assertion.

we have worked so long on such a problem and still lack the standardization that the World Wide Web, for example, seems to have acquired overnight? Some researchers within our community assert that at least part of the problem lies in a failure to collaborate effectively. In his opening address to the International Medical Informatics Association Working Group 6 (IMIA WG6) Conference on Natural Language and Medical Concept Representation, Jean-Raoul Scherrer summarized his dismay with the following statement:

> We are faced with overwhelming new problems! Indeed, there are now "doctrinarian chapels" that are reluctant to collaborate with one another, opposed to other open, pragmatic groups finding their way, transgressing the doctrines, not always consciously and hence without having the least guilty feeling whatsoever. Let us now enter into a more reflective phase of much needed psychotherapy.[2]

Scherrer also suggests in his opening address that we have been encumbered with jargon that reinforces separation by noting that "domains" in database theory, "types" in artificial intelligence, "classes" in object-oriented approaches, and "sorts of" in predicate calculus are all different names given for the representation of "categories of beings."

Scherrer provides a clear voice for reconciliation and consolidation from within the medical-informatics community, but there have been similar appeals from the knowledge-representation community as well. Schubert,[3] for example, asks that community to consolidate ideas, discarding artificial distinctions and inconsistent terminology. He argues that all knowledge-representation schemes that aspire to cope with large, general, propositional knowledge bases, such as "frame-based systems," "semantic databases," "blackboard systems," and "semantic networks" are in fact just a set of first-order predicate-calculus formulas, in a "slightly altered terminological guise," and therefore are examples of artificial distinctions derived from individuals who find it "faster to rediscover something within the framework of one's colony than to glean it from the writings of another." Schubert[3] supports his claim of the equivalence of various methods for representing knowledge bases with a formal proof. However, he does not provide us with a framework upon which we can build in our effort to consolidate ideas in those settings where we have shared understanding sufficient to reach a collaborative consensus, while we struggle to identify terminologic relationships and to develop consensus in other areas where our understanding remains limited. Perhaps the problem results not from a failure to collaborate, but rather from a failure to collaborate at the right level of abstraction—a level at which we have suffi-

cient collective experience and understanding to reach consensus.

If we are to consolidate our ideas, where should we begin? A pragmatic first step is to define clearly the goals and limitations of a collaborative activity. Early UMLS collaborators worked for many years to define the boundaries embodied within the UMLS Metathesaurus.[4] We believe that this embodiment has served the UMLS project well in that it encourages terminology developers to explore diverse semantic models while allowing those terminologies to be integrated via the Metathesaurus without requiring reconciliation of possibly orthogonal semantic models.

The Metathesaurus accomplishes this integration through the Concept Unique Identifier (CUI), a conceptual "key" that takes on emergent meaning through its linkage to concepts in different source terminologies that are in some sense equivalent.† As such, the Metathesaurus provides a framework where researchers with divergent perspectives can collaborate at a high level, while maintaining the freedom to pursue independent ideas (which may be reconciled after they have demonstrated their viability and consensus is developed regarding their merits). By allowing high-level collaboration without forcing reconciliation of orthogonal semantic models, the UMLS can achieve its stated goal "to facilitate the development of conceptual connections between users and relevant machine-readable information"[5] without waiting for a consensus about deeper semantic issues that may never be realized.

To maintain an effective collaboration, ongoing evaluations of the collaborative products, processes, and objectives is vital. The UMLS has been the focus of many evaluations done by researchers directly funded by the National Library of Medicine as well as by others who have chosen to collaborate with groups that are involved with the UMLS project. These evaluations have helped to shape our collective understanding of the UMLS and in some cases have sought to ascertain how the UMLS may perform tasks other than its stated objective of developing conceptual connections between users and machine-readable information.

---

†Equivalence is of course relative to some perspective. In the Metathesaurus for example, aspirin, Ecotrin (a formulation of aspirin with an enteric coating), and Aspergum (a formulation of aspirin in a chewing-gum base) are considered equivalent (these three aspirin-related concepts exist as distinct concepts in at least one of the UMLS sources, but are linked to the same CUI and thus determined equivalent—presumably in the sense that they share the same active ingredients). From other perspectives—such as information systems that record medications dispensed to patients—aspirin, Ecotrin, and Aspergum would not be considered equivalent.

Evaluation of the UMLS

The UMLS is an applied work—it seeks to develop content and applications for a stated objective. As such, there are two obvious areas for evaluation: (1) the quality of the UMLS content, and (2) the ability of applications to provide conceptual connections between users and machine-readable information using UMLS content. It would also be relevant to examine the quality and effectiveness of the collaborative activities underlying the development of the UMLS.

Evaluation of Content

Medical terminologies include phrases used in our language and codify them in an effort to facilitate communication of medical thoughts in a reproducible way. The ability of these terminologies to cover relevant concepts in a particular domain has been described as the "content coverage" of the terminology. There have been many projects to evaluate the content coverage of the UMLS. Researchers conducting these studies have typically taken clinical notes from medical records, manually extracted clinically relevant phrases, and looked up those phrases in the UMLS to see if equivalent concepts exist. Specific areas of clinical content that have been analyzed include: the subjective, objective, assessment, and plan portions of clinic notes for terms related to hypertension,[6] chest x-ray reports,[7] nurses' notes,[8] ambulatory family-practice records,[9] history and physical dictations, nursing notes, consulting notes, outpatient progress notes, inpatient progress notes, discharge summaries, radiology reports, operative reports,[10,11] problem lists from primary care sites,[12] clinical laboratory terms, including laboratory tests and panels, measured substances, and sampled substances,[13] genetic disorders,[14] and North American Nursing Diagnosis List of Approved Diagnoses and nursing problem terms and intervention terms.[15]

In several of these studies, the investigators have "scored" the UMLS in some respect against the very source terminologies from which it is constructed.[6,10,11,13,15,16] Many readers of these studies have used such scores as metrics for determining what the "best" terminology system would be to represent medical records—even though the authors of these studies often clearly state that the UMLS is not itself a system for primary representation of patient data.

We agree that looking for the coverage of typical medical phrases is a useful evaluation metric for the UMLS, since it is from such measures that we can determine if the coverage of medical language by the UMLS is improving and can thereby prioritize future work. We argue, however, that scoring the coverage of the UMLS *as if it were a competitor* for the task of encoding medical records is not useful, and in fact propagates a misunderstanding about the purpose of the UMLS. This common misunderstanding can be minimized if future evaluations can define ways to characterize the synergistic relationship between the source terminologies upon which the UMLS is founded, and the value that the UMLS provides by integrating these sources into a common framework.

Many researchers have proposed desiderata for terminology systems.[17,18] By re-examining those criteria with the explicit goal of differentiating the value provided by a source terminology, and the value provided by integrating those terminologies into a common framework, we can determine when metrics should be applied to a source or to the UMLS, and we can also provide a context for interpreting those results. Although development of a complete set of metrics for evaluating terminology systems is beyond the scope of this article, we believe that efforts toward articulating appropriate distinctions can serve those who seek to quantify the content coverage and the concept organization of terminology systems.

Coverage of medically relevant phrases is an important property of the UMLS, although this property is derived from the phrases contained within the sources it incorporates. A property unique to the UMLS Metathesaurus is the integration of these sources into a framework within which equivalent and related entities can be linked to one another as well as to the semantic neighborhood that results from this linkage. By objectively evaluating the comprehensiveness of the linkages and the properties of the semantic neighborhood, we can learn how the UMLS should be modified to progress towards its goal of linking concepts' machine-readable representations in ways that reflect how they are understood by human beings. Cimino describes methods for evaluating such UMLS properties.[19]

Evaluation of Functionality

The National Research Council reported on methods for evaluating applied works in computer science.[20] In their report, they articulated a framework that can be validly applied to the evaluation of UMLS. They describe three tiers of evaluation that should follow research projects from their inception to their successful conclusion. The first is *proof of existence*: demonstration of a fundamentally new method (an example would be the conception and demonstration of functionality of the initial computer "mouse" as a method for human–computer interaction). The second is *proof of concept*: demonstration that a method can be applied in some circumstances and perform appropriately. The third is *proof of performance*: demonstration that a method can be applied routinely and can perform better than competing alternatives.

For the UMLS, the past decade has contained many such demonstrations, beginning with the construction of Meta-0 (a proof of existence for the Metathesaurus),[21] and continuing with the demonstration of applications that successfully link clinical systems with knowledge-based information sources (proofs of concept for developing conceptual connections between users and machine-readable information).[22–28] In the next decade, as the NLM makes the transition from its experimental UMLS versions (last released in 1996) to annual UMLS versions (first released in 1997), we expect that evaluation of the functionality enabled by the UMLS will be ongoing and will meet either meet the standards of a proof of performance—demonstration of the routine applicability and superior performance of the UMLS—or will prompt evolution of the UMLS such that high standards are achieved and maintained.

Evaluation of Collaboration

A remarkable number of papers have been written regarding the UMLS, its evaluation, its impact, and its potential role in biomedical information systems.‡ Although the UMLS is undoubtedly a result of a great deal of cooperation and distributed effort, it is less clear whether a true collaboration has developed among the various groups who maintain the component vocabularies from which the UMLS is constructed. We are unaware of any specific evaluations of this issue. Smith[29] has argued that there is more to collaboration than two or more individuals jointly working on a project: "Collaboration carries with it the expectation of a singular purpose and a seamless integration of parts, as if the conceptual object were produced by a single good mind. A requirement for collective intelligence is achieving a critical level of coherence in the work of the group." Cooperative efforts are largely equal to the sum of their parts, whereas collaborative efforts are more synergistic and interdependent.[30]

The UMLS would have much to gain from a more clearly collaborative effort in which shared models were more aggressively sought. We acknowledged earlier that the UMLS provides a framework within which researchers and developers can reconcile common aspects of their approaches, while ensuring that there is an opportunity within this framework to explore unique approaches and to meet specialized needs. However, it is also likely that many of the terminologic differences that appear among component vocabularies could be reconciled if there were a more explicit effort for the maintainers of the UMLS's

source terminologies to work together to agree on more consistent representations for shared concepts.

Summary

We believe it is through ongoing collaboration and evaluation that the UMLS will grow in new directions that better support the challenges of mapping between terminologies that address overlapping aspects of health-care, data-representation needs.

The past decade of UMLS work has focused pragmatically upon developing an integrating framework using applications to derive structure from the language of source terminology systems. The work has also sought to capture any structure provided formally by those systems. The next decade will see significant advances in the integration of the diverse approaches to the terminology of the UMLS if the informatics community understands and leverages the unifying paradigm that the UMLS offers, and if the UMLS continues to grow by directly supporting increasingly formal terminology systems. Through such extension, the UMLS will attract a broader range of collaborators, provide new functionality, and continue to evolve to meet the needs of the informatics community (and, in turn, those of the end users of our systems) more effectively. Furthermore, the community will better understand the role of the UMLS as a terminology mapping, translation, and maintenance system on which broad-based information retrieval and domain modeling can be based.

By unifying terminologies that serve different tasks for different disciplines, we define how aspects of our disciplines can be unified as well. This unification may turn out to be one of the UMLS's greatest legacies, but such an outcome will depend on the contributions and commitments of the informatics profession.

Our profession cannot reconcile and consolidate ideas and approaches to standardization of terminology unless we assure that the UMLS becomes a living organism that adapts to current thinking in medicine and in informatics. As a profession, we can do so by making the UMLS our own, working closely with the National Library of Medicine and those who maintain and adapt this important national resource. Our profession must promote communication among users of the UMLS, developers of the source terminologies, developers of software systems that use the UMLS and its source terminologies, and the developers of the UMLS itself. The Large Scale Vocabulary Test is one such effort to do this.[31] We accordingly urge and anticipate many more such efforts initiated both by both the National Library of Medicine and the informatics community at large.

References ■

1. Sittig DG. Grand challenges in medical informatics? J Am Med Inform Assoc. 1994;1:412–13.

2. Scherrer J-R. Concepts, knowledge and language in health-care information systems: follow-up 30 months later. In: Chute CG (ed). Conference on Natural Language and Medical Concept Representation. Jacksonville, Florida, 1997;5–8. (Also to appear in a special issue of Methods of Information in Medicine.)

3. Schubert LK. Semantic nets are in the eye of the beholder. In: Sowa JF (ed). Principles of Semantic Networks: Explorations in the Representation of Knowledge. San Mateo: Morgan Kaufmann, 1991;95–107. (Brachman RJ, ed. Morgan Kaufmann Series in Representation and Reasoning.)

4. Humphreys BL, Lindberg DAB, Schoolman HM, Barnet OG. The Unified Medical Language System: An informatics research collaboration. J Am Med Inform Assoc. 1998;1:1–11.

5. Humphreys BL, Lindberg DAB. The UMLS Project: Making the conceptual connection between users and the information they need. Bulletin of the Medical Library Association 1993;81:170–77.

6. Campbell JR, Kallenberg GA, Sherrick RC. The clinical utility of META: an analysis for hypertension. In: Frisse ME (ed). Sixteenth Annual Symposium on Computer Applications in Medical Care. New York: McGraw-Hill, 1992;397–401.

7. Friedman C. The UMLS coverage of clinical radiology. In: Frisse ME (ed). Sixteenth Annual Symposium on Computer Applications in Medical Care. New York: McGraw-Hill, 1992;309–13.

8. Lange LL. Representation of everyday clinical nursing language in UMLS and SNOMED. In: Cimino JJ (ed). AMIA Fall Symposium. Philadelphia: Hanley & Belfus, 1996;140–44.

9. Mullins HC, Scanland PM, Collins D, et al. The efficacy of SNOMED, Read Codes, and UMLS in coding ambulatory family practice clinical records. In: Cimino JJ (ed). AMIA Fall Symposium. Philadelphia: Hanley & Belfus, 1996;135–39.

10. Chute CG, Cohn S, Campbell KE, Oliver D, Campbell JR. The content coverage of clinical classifications. J Am Med Inform Assoc. 1996;3:224–33.

11. Campbell JR, Carpenter P, Sneiderman C, Cohn S, Chute CG, Warren J. Phase II evaluation of clinical coding schemes: Completeness, taxonomy, mapping, definitions, and clarity. J Am Med Inform Assoc. 1997;4:238–51.

12. Campbell JR, Payne TH. A comparison of four schemes for codification of problem lists. In: Osbolt JG (ed). Eighteenth Annual Symposium on Computer Applications in Medical Care. Philadelphia: Hanley & Belfus, 1994;201–5.

13. Cimino JJ. Representation of clinical laboratory terminology in the Unified Medical Language System. In: Clayton PD (ed). Fifteenth Annual Symposium on Computer Applications in Medical Care. New York: McGraw-Hill, 1991;199–203.

14. O'Keefe K, Sievert ME, Mitchell JA. Mendelian Inheritance in Man: Diagnoses in the UMLS. In: Safran C (ed). Seventeenth Annual Symposium on Computer Applications in Medical Care. New York: McGraw-Hill, 1993;735–39.

15. Zielstorff RD, Cimino C, Barnet GO, Hassan L, Blewett DR. Representation of nursing terminology in the UMLS Metathesaurus: A pilot study. In: Frisse ME (ed). Sixteenth An-nual Symposium on Computer Applications in Medical Care. New York: McGraw-Hill, 1992;392–96.

16. Friedman C, Cimino JJ, Johnson SB. A schema for representing medical language applied to clinical radiology. J Am Med Inform Assoc. 1994;1:233–48.

17. Cimino JJ, Hripcsak G, Johnson SB, Clayton PD. Designing an introspective, controlled medical vocabulary. In: Kingsland LW (ed). Thirteenth Annual Symposium on Computer Applications in Medical Care. Washington, DC: IEEE Computer Society Press, 1989;513–18.

18. Evans DA, Cimino JJ, Hersh WR, Huff SM, Bell DS. Toward a medical-concept representation language. J Am Med Inform Assoc. 1994;1:207–17.

19. Cimino JJ. Auditing the Unified Medical Language System with Semantic Methods. J Am Med Inform Assoc. 1998;5:41–51.

20. National Research Council. Academic careers for experimental computer scientists and engineers. Washington, DC. National Adademy Press, 1994.

21. Sheretz DD, Tuttle MS, Olson NE, Erlbaum MS, Nelson SJ. Lexical mapping in the UMLS Metathesaurus. In: Kingsland LW (ed). Thirteenth Annual Symposium on Computer Applications in Medical Care. Washington, DC: IEEE Computer Society Press, 1989;494–9.

22. Cimino JJ. Linking patient information systems to bibliographic resources. Methods of Information in Medicine. 1996;35:122–6.

23. Lowe HJ, Walker WK, Polonkey S, Jiang F, Vries JK, McCray AT. The image engine HPCC project: A medical digital library system using agent-based technology to create an integrated view of the electronic medical record. In: Adam N, Halem M, Yesha Y (eds). Third Forum on Research and Technology Advances in Digital Libraries, ADL '96. Los Alamitos: IEEE Computer Society Press, 1996;45–56.

24. Miller PL, Frawley SJ. Trade-offs in producing patient-specific recommendations from a computer-based clinical guideline: a case study. J Am Med Inform Assoc. 1995;2:238–42.

25. Miller RA, Gieszczykiewicz FM, Vries JK, Cooper GF. CHARTLINE: providing bibliographic references relevant to patient charts using the UMLS Metathesaurus Knowledge Sources. In: Frisse ME (ed). Sixteenth Annual Symposium on Computer Applications in Medical Care. New York: McGraw-Hill, 1992;86–90.

26. Nelson SJ, Sherertz DD, Tuttle MS. Issues in the development of an information retrieval system: The Physician's Information Assistant. In: Lun KC, Degoulet P, Piemme TE, Rienhoff O (eds). MEDINFO 92. Geneva, Switzerland: North Holland: Elsevier Science Publishers, 1992;371–5.

27. Powsner SM, Miller PL. Automated online transition from the medical record to the psychiatric literature. Meth Inform Med. 1992;31:169–74.

28. Mulligen EM. UMLS-based access to CPR data. In: Cimino JJ (ed). AMIA Fall Symposium. Philadelphia: Hanley & Belfus, 1996;819.

29. Smith J. Collective intelligence in computer-based collaboration. Hillsdale, NJ: Lawrence Erlbaum, 1993.

30. Shortliffe E, Patel V, Cimino J, Barnet G, Greenes R. A study of collaboration among medical informatics research laboratories. AI in Medicine. 1998; (in press).

31. Humphreys BL, McCray AT, Cheh ML. Evaluating the coverage of controlled health data terminologies: report on the results of the NLM/AHCPR Large Scale Vocabulary Test. J Am Med Inform Assoc. 1997;4:483–497.