MediGrid – Facilitating Semantic-Based processing of Biomedical Data and Knowledge

Jan Vejvalka[1], Petr Lesný[1], Tomáš Holeček[2], Kryštof Slabý[1], Adéla Jarolímková[3] and Helena Bouzková[4]

[1]Faculty hospital in Prague Motol
V Úvalu 84, 150 06, Prague 5 – Motol, Czech Republic

[2]FMS FHS Charles University
U Kříže 10, 158 00 Praha 5 Jinonice, Czech Republic

[3]CESNET, z.s.p.o.
Zikova 4, 160 00 Praha 6, Czech Republic

[4]National Medical Library
Sokolská 54, 121 32 Praha 2, Czech Republic

Abstract. New ICT approaches promise better support to the traditional ways that medicine handles information. Exploring ways to provide this support, we formulated the principles of he MediGrid: (1) Data processed by biomedical algorithms are (following the philosophical tradition of phenomenology) indicators that can be transformed into other indicators. (2) Data and algorithms can be shared across conceptual domains if documented semantic links exist to support such interconnection. The need of explicit and detailed documentation of semantics leads to the requirement of good documentation of computer procedures that implement the biomedical knowledge contained in scientifically accepted algorithms. A proof-of-concept implementation of a system based on these principles has been published on Sourceforge.

1 Introduction

Fast growing volume of scientific knowledge and resulting specialization in medicine stresses the utilization of algorithms and guidelines, that objectify and standardize healthcare processes. The idea of using ICT to support these algorithms and guidelines is natural: execution of algorithms is what information technologies usually do. Algorithms of care, however, are executed on a different level, and with different data than computer algorithms. Proper use of ICT requires a proper computer representation of these algorithms and data. Several projects aiming at creating or documenting a collection of biomedical algorithms exist (such as MEDAL, MedCalc etc.).

Vejvalka J., Lesný P., Holeček T., Slabý K., Jarolímková A. and Bouzková H. (2009).
MediGrid – Facilitating Semantic-Based processing of Biomedical Data and Knowledge.
In *Proceedings of the 1st International Workshop on Open Source in European Health Care: The Time is Ripe*, pages 18-21
DOI: 10.5220/0001828500180021

A problem that has been complicating ICT support to algorithms in healthcare is the nature of information that healthcare deals with, its non-trivial semantics and traditional models of handling this information. To make automatic computerized processing of biomedical information possible, its semantic value must be properly represented (and evaluated) in a computer-comprehensible form. To facilitate correct processing of biomedical data and their semantics, we have designed the MediGrid, a working tool for representation and processing of domain ontologies in biomedicine.

2 Materials and Methods

MediGrid aims to facilitate automatic processing of biomedical data with scientifically relevant knowledge transformed into computational algorithms, using up-to-date information technologies.
MediGrid approach is based on representation of biomedical data and algorithms as resources, and on sharing these resources in a grid-like environment. The natural way of assigning biomedical algorithms to data (or vice versa) is by comparing their semantic values; therefore MediGrid builds its mechanisms of sharing on capturing and processing of semantic information. Results from previous projects (SMARTIE, Growth) have been used as a base for our work: we used existing algorithms (Growth, MEDAL) to verify our results and we used our experience from SMARTIE to structure semantic information.

3 Results

Based on extensive research and on previous work, we have postulated the basis of MediGrid as the working tool for analyzing algorithm-related ontologies in medicine and for realizing (and practically implementing) their potential for data processing. The key principles of MediGrid can be expressed as follows:

1. Data processed by biomedical algorithms are (following the philosophical tradition of husserlian phenomenology) indicators that can be transformed into other indicators and grouped into indicator classes by their roles in these transformations.
2. Data and algorithms can be shared across conceptual domains if documented semantic links exist to support such interconnection.

The emphasize on semantics of both algorithms and data, the need for its extensive review and verification when linking data with algorithms, and also the stress on correct procedures (lege artis) in medicine bring practical requirements on design and implementation of a knowledge-processing system based on these theoretical principles:

- Semantic information (meaning for the human user) of both indicator classes and transformations must be explicitly described and available for assessment and validation.

- Semantic information must be bound to the current scientific paradigm and to evidence based medicine through extensive links to published and reviewed works.
- Mechanisms of procedural authority and trust must be implemented to support users' decisions about procedural and semantic values of individual components.

We have identified several entities, which have to be documented in order to achieve the aforementioned requirements. These entities can be set up in 4 layers:

- The source layer, which contains the description of MediGrid information: author and cited work.
- The algorithm layer, which contains the description of the most commonly used categories – the transformation (algorithm), indicator class (the data entered or exchanged between the algorithms)
- The implementation layer, which contains information about the specific implementation of transformations and of validations of indicator classes in computer programs.
- The review/trust layer, containing the user reviews and trust statements.

Description of each entity contains four basic elements:

- human semantics – the collection of human readable description, which enables user to understand the meaning of the entity
- metadata – pieces of computer recognizable data, which can be utilized e.g. to construct the user interfaces
- relations – relations of this entity to other entities
- classifications – special case of relations that position the entity in external classification systems and ontologies (UMLS, MeSH descriptors etc.).

Based on the described principles, we posted our first (proof of concept) open source implementation of MediGrid to Sourceforge: http://medigrid.sourceforge.net.

4 Conclusions

The need for explicit documentation of biomedical data and algorithms and their semantics on several layers leads to the requirement of good documentation of computer procedures that implement the biomedical knowledge contained in scientifically accepted algorithms. Similarly, in order to be open to scientific scrutiny, also the framework system for processing of biomedical information has to be sufficiently documented at all levels.

Acknowledgements

Supported by Czech research projects MediGrid, 1ET202090537 and by VZ FNM, MZO 00064203.

References

1. Krásničanová H., Lesný P.: Compendium of pediatric auxology 2005. Praha, Novo Nordisk (2005).
2. Svirbely J, Iyengar MS.: Issues in the implementation of computer-based medical algorithms. Technology and Health Care. 2005; 13(5); 438-439.
3. Smart Medical Applications Repository of Tools For Informed Expert (S.M.A.R.T.I.E.) and MedNotes™. IST (2000). http://www.smartie-ist.org
4. Lipscomb C.E.: Medical Subject Headings (MeSH). Bull Med Libr Assoc (2000) 265-6.
5. Cote R.A., Robboy S.: Progress in medical information management. Systematized nomenclature of medicine (SNOMED). Jama (1980) 756-62.
6. Lindberg C.: The Unified Medical Language System (UMLS) of the National Library of Medicine. J Am Med Rec Assoc (1990) 40-2.
7. Husserl E. Logical Investigations, Investigation I. (translated by J. N. Finday). Routledge, London (2001) 184.
8. Kuba M., Krajíček O., Lesný P., Vejvalka J., Holeček T.: Grid empowered sharing of medical expertise. Stud Health Technol Inform (2006) 273-82.
9. Bodenreider O.: Biomedical ontologies in action: role in knowledge management, data integration and decision support. Yearb Med Inform. 2008:67-79.