

One Empirical Study, Three Tests of Translation, Many Questions on Biomedical Ontologies: Limited Contribution of MeSH Terms to Effective Literature Searches on ‘Health-Related Values’

Mila Petrova

Egenis (ESRC Centre for Genomics in Society), University of Exeter, Devon, UK

Keywords: biomedical ontologies, ontologies of values, values in ontologies, search strategies, research-into-practice translation, evidence-based medicine, bio-psycho-social frameworks, boundary objects, promiscuous realism

This paper will raise three translation issues associated with biomedical ontologies: translation between biomedical ontologies and health-related ontologies informed by the social and behavioral sciences and the humanities; translation of ontologies into literature searching strategies in electronic bibliographic databases; and translation of intended applications of ontologies to the indexing of research publications into actual indexing practices.

The broad context in which these issues have been articulated is that of literature searches underpinning ‘research synthesis’ work (of which systematic reviews are the most influential type, but at least forty other methods have been identified in review publications). The narrow context is that of a study on developing search strategies for identifying publications on health-related values in electronic bibliographic databases. ‘Values’ was understood very broadly, to include issues such as patients’ and other stakeholders’ perceptions, preferences, beliefs, experiences, expectations, quality of life; issues of inter-disciplinary communication; values underpinning diagnostic classifications, etc. The core element of the study was a word frequency analysis of datasets on Diabetes, Obesity, Dementia and Schizophrenia. These comprised 4,440 citations (2,449 “true positives” and 1,991 “false positives” for values contents; MEDLINE; Jan 2004 – Dec 2006) amounting to over a million-word textual corpus. Both text words and MeSH terms were analyzed. A 22-line values search filter was developed following principles of objective methods for search filter development. It had sensitivity and precision of 76.8% and 86.8% in the development dataset and between 47.1% and 70.1% (sensitivity) and 63.6% ÷ 82.6% (precision) in validation datasets. Text words came out as significantly more effective than MeSH terms relative to the desiderata for the filter (brief, of high precision and at least moderate sensitivity).

The work also showed unexpected patterns of assignment of terms to research papers for the purposes of MEDLINE indexing. For instance, it seems that “Attitude to Health” or “Health Knowledge, Attitudes, Practice” act as umbrella terms for almost anything “psychosocial”.

The first translation issue this study pointed to concerns the relationship between biomedical ontologies on the one hand and health-related ontologies coming from the social and behavioral sciences and the humanities on the other. Ontologies underpinning medical and health bibliographic databases were found to be either minimal in mapping the field of health-related values or aligned with concepts from the social and behavioral sciences and the humanities that do not translate easily into clinical practice and/or health policy concerns. Can the biomedical ontologies community justifiably treat this as an NMP (“not my problem”)? Could engagement with those “other” health-related ontologies enhance a biomedical ontology’s own capacity to support the translation of research into clinical practice?

The second translation issue identified concerns the capacity of biomedical and other health-related ontologies to contribute to effective literature searching strategies. In this study, controlled vocabulary terms had a limited role in an objectively developed search filter. Text words comprised by far the greater proportion of it (18 vs. 3 terms). Thus, current ontologies mapping the field of health-related values failed to translate into effective literature searching strategies. Is this a generic problem, of difficulties of translation between highly discriminative, fine-grained ontologies and coarser-grained user questions? Or is it a middle-ground one, for instance of translation of some ontologies into literature searches on broad, fuzzy topics? Or is it simply a very particular problem solvable through established means?

The third translation issue highlighted by this study concerns the actual, practically negotiated use of ontologies in the indexing of research publications. Findings suggested that certain patterns of assignment of controlled vocabulary terms to research papers may reflect historical, local and contingent practices rather than a reasonably direct matching of the explicit definition and scope of a term and the contents of a research paper. How can knowledge of such actual 'workarounds' help us improve the contents and structure of our ontologies? Are current feedback mechanisms between ontology developers and

different types of ontology users robust and varied enough?

To conclude, I will discuss whether the development and application of ontologies supporting research synthesis work or mapping fields of fuzzy concepts such as health-related values may be usefully informed by Susan Leigh Star's concept of "boundary objects" and John Dupré's ideas on pluralism and "promiscuous realism". I will also suggest some priorities for empirical research on the actual use of biomedical ontologies in terms of translation issues.