World Academy of Science, Engineering and Technology
International Journal of Bioengineering and Life Sciences
Vol:17, No:01, 2023

Linguistic Insights Improve Semantic Technology in Medical Research and Patient Self-Management Contexts

Authors : William Michael Short

Abstract : Semantic Web' technologies such as the Unified Medical Language System Metathesaurus, SNOMED-CT, and MeSH have been touted as transformational for the way users access online medical and health information, enabling both the automated analysis of natural-language data and the integration of heterogeneous healthrelated resources distributed across the Internet through the use of standardized terminologies that capture concepts and relationships between concepts that are expressed differently across datasets. However, the approaches that have so far characterized 'semantic bioinformatics' have not yet fulfilled the promise of the Semantic Web for medical and health information retrieval applications. This paper argues within the perspective of cognitive linguistics and cognitive anthropology that four features of human meaning-making must be taken into account before the potential of semantic technologies can be realized for this domain. First, many semantic technologies operate exclusively at the level of the word. However, texts convey meanings in ways beyond lexical semantics. For example, transitivity patterns (distributions of active or passive voice) and modality patterns (configurations of modal constituents like may, might, could, would, should) convey experiential and epistemic meanings that are not captured by single words. Language users also naturally associate stretches of text with discrete meanings, so that whole sentences can be ascribed senses similar to the senses of words (so-called 'discourse topics'). Second, natural language processing systems tend to operate according to the principle of 'one token, one tag'. For instance, occurrences of the word sound must be disambiguated for part of speech: in context, is sound a noun or a verb or an adjective? In syntactic analysis, deterministic annotation methods may be acceptable. But because natural language utterances are typically characterized by polyvalency and ambiguities of all kinds (including intentional ambiguities), such methods leave the meanings of texts highly impoverished. Third, ontologies tend to be disconnected from everyday language use and so struggle in cases where single concepts are captured through complex lexicalizations that involve profile shifts or other embodied representations. More problematically, concept graphs tend to capture 'expert' technical models rather than 'folk' models of knowledge and so may not match users' common-sense intuitions about the organization of concepts in prototypical structures rather than Aristotelian categories. Fourth, and finally, most ontologies do not recognize the pervasively figurative character of human language. However, since the time of Galen the widespread use of metaphor in the linguistic usage of both medical professionals and lay persons has been recognized. In particular, metaphor is a well-documented linguistic tool for communicating experiences of pain. Because semantic medical knowledge-bases are designed to help capture variations within technical vocabularies – rather than the kinds of conventionalized figurative semantics that practitioners as well as patients actually utilize in clinical description and diagnosis – they fail to capture this dimension of linguistic usage. The failure of semantic technologies in these respects degrades the efficiency and efficacy not only of medical research, where information retrieval inefficiencies can lead to direct financial costs to organizations, but also of care provision, especially in contexts of patients' self-management of complex medical conditions.

Keywords : ambiguity, bioinformatics, language, meaning, metaphor, ontology, semantic web, semantics
Conference Title : ICBNLP 2023 : International Conference on Biomedical Natural Language Processing
Conference Location : London, United Kingdom
Conference Dates : January 23-24, 2023