

CiTO + SWAN: The Web Semantics of Bibliographic Records, Citations, Evidence and Discourse Relationships

Paolo Ciccarese^{a,b *}, David Shotton^{c *}, Silvio Peroni^{c,d} and Tim Clark^{a,b §}

^a Massachusetts General Hospital, Department of Neurology, Mindinformatics, 65 Landsdowne Street, Cambridge, 02139 MA, USA.

^b Harvard Medical School, Boston, MA, USA.

^c Image Bioinformatics Research Group, Department of Zoology, University of Oxford, South Parks Road, Oxford OX1 3PS, UK.

^d Department of Computer Science, University of Bologna, Mura Anteo Zamboni 7, 40127 Bologna (BO), Italy

* These authors contributed equally to this work.

§ Corresponding author. Email: tim_clark@harvard.edu

Abstract. Most literature searching in biomedicine is now conducted via PubMed, Google Scholar or other web-based bibliographic search mechanisms. Yet until now a public, open, interoperable and complete web-adapted information schema for bibliographic citations, bibliographic references and scientific discourse has not been available. Such a schema, expressed in the form of a description logic compatible with current web semantics approaches, would provide the ability to treat bibliographic references and citations, and rhetorical discourse in scientific publications, as semantic metadata on the web, with all the benefits that implies for organization, search and mash-up of web-based scientific information.

In this paper we present CiTO + SWAN, a set of fully harmonized ontology modules resulting from the harmonization of CiTO (the Citation Typing Ontology) with SWAN (Semantic Web Applications in Neuromedicine), which we have developed by jointly adapting and evolving version 1.6 of CiTO, the Citation Typing Ontology, and version 1.2 of the SWAN Scientific Discourse Ontology (v1.2). The CiTO + SWAN model is specified in OWL 2 DL, is fully modular, and inherently supports agent-based searching and mash-ups.

Through the harmonization activity presented here, and previous work that harmonized SWAN with the SIOC (Semantically-Interlinked Online Communities) Ontology for describing blogs, wikis and discussion groups, we have construct the basis of a powerful new web framework for scientific communications.

Keywords: Bibliographic ontology, FRBR, scientific discourse, OWL

