FaBiO and CiTO: ontologies for describing bibliographic resources and citations

Silvio Peroni
Department of Computer Science, University of Bologna, Mura Anteo Zamboni 7, 40126 Bologna, Italy.
essepuntato@cs.unibo.it

David Shotton
Research Data Management and Semantic Publishing Research Group,
Department of Zoology, University of Oxford
South Parks Road, Oxford OX1 3PS, UK.
david.shotton@zoo.ox.ac.uk

Keywords: OWL, publication characterization, scholarly reference, semantic publishing

Abstract

Semantic publishing is the use of Web and Semantic Web technologies to enhance the meaning of a published journal article, to facilitate its automated discovery, to enable its linking to semantically related articles, to provide access to data within the article in actionable form, and to facilitate integration of data between articles. Recently, semantic publishing has opened the possibility of a major step forward in the digital publishing world. For this to succeed, new semantic models and visualization tools are required to fully meet the specific needs of authors and publishers. In this article, we introduce the principles and architectures of two new ontologies central to the task of semantic publishing: FaBiO, the FRBR-aligned Bibliographic Ontology, an ontology for recording and publishing bibliographic records of scholarly endeavours on the Semantic Web, and CiTO, the Citation Typing Ontology, an ontology for the characterization of bibliographic citations both factually and rhetorically. We present those two models step by step, in order to emphasise their features and to stress their advantages relative to other pre-existing information models. Finally, we review the uptake of FaBiO and CiTO within the academic and publishing communities.

1 Introduction

Scholarly authoring and publishing are in the throes of a revolution, as the full potential of on-line publishing is explored. Yet, to date, publishers have not adopted Web standards for their work, but rather employ a variety of proprietary XML-based informational models and document type definitions (DTDs). While such independence was reasonable in the pre-web world of paper publishing, it now appears anachronistic, since publications and their metadata from different sources are incompatible, requiring hand-crafted mappings to convert from one to another. For a large community such as publishers, this lack of standard definitions that could be adopted and reused across the entire industry represents losses in terms of money, time and effort.

In contrast, modern web information management techniques employ standards such as RDF [4] and OWL 2 [28] to encode information in ways that permit computers to query metadata and integrate web-based information from multiple resources in an automated manner. Since the processes of scholarly communication are central to the practice of science, it is essential that publishers now adopt such standards to permit inference over the entire corpus of scholarly communication represented in journals, books and conference proceedings. This requires the availability of appropriate ontologies that are specially tailored to the requirements of authors, publishers and their readers. The purpose of this paper is to present two such ontologies that form key components of the semantic publishing revolution.

Semantic publishing is the use of web and semantic web technologies to enhance a published document such as a journal article so as to enrich its meaning, to facilitate its automatic discovery, to enable its linking to semantically related articles, to provide access to data within the article in actionable form, and to allow integration of data between papers [21] [24]. Semantic publishing and scholarly citation using web standards are presently two of the most interesting topics within the scientific publishing domain. Research areas in this domain include the development of:

- *semantic models* (vocabularies, ontologies) that meet the requirements of scholarly authoring and publishing;
- *visualization and documentation tools* that permit such ontologies to be easily understood;
- *annotation tools* that allow these models to be used for enhancing documents with relevant semantic assertions;
- *new algorithms* to take advantages of these semantic annotations when searching over large sets of on-line documents.

In this article, we address the first point by describing the principles and architecture of two ontologies central to the task of semantic publishing: **FaBiO**, the *FRBR-aligned Bibliographic Ontology*[1], an ontology for recording and publishing bibliographic records of scholarly endeavours on the Semantic Web, and **CiTO**, the *Citation Typing Ontology*[2], an ontology for the characterization of bibliographic citations, both factually and rhetorically. These ontologies are members of SPAR, the *Semantic Publishing and Referencing Ontologies*[3], a suite of orthogonal and complementary OWL 2 DL ontology modules that together permit the creation of comprehensive machine-readable RDF metadata for all aspects of semantic publishing and referencing.

The rest of this article is organized as follows: in Section 2 we introduce principles that have guided our work, and in Section 3 we briefly describe the existing ontologies and vocabularies we took into account when developing FaBiO

1 FaBiO, the FRBR-aligned bibliographic ontology: http://purl.org/spar/fabio.
2 CiTO, the Citation Typing Ontology: http://purl.org/spar/cito.
3 The SPAR Ontologies: http://purl.org/spar.

and CiTO. These new ontologies are then presented in Section 4 and Section 5 respectively. In Section 6 we describe the contexts in which FaBiO and CiTO are now being employed, and finally in Section 7 we sketch out some conclusions and future directions of our work.

2 Characteristics, starting point and principles

The main characteristics of this work, that mark it out as distinct from previous contributions, is the creation of two new semantic publishing and referencing ontologies of sufficient expressivity to meet the requirements of end users such as academic authors and publishers.

We have also developed two new presentation technologies, the *Live OWL Documentation Environment* (*LODE*)[4] and the *Graphical Framework For OWL Ontologies* (*Graffoo*)[5], which support the interaction between end users and ontologies. LODE can be seen in action by opening the URL http://purl.org/spar/fabio/ in a Web browser window (and is described in [18]), while the use of Graffoo is exemplified in Figures 1 and 2 of this paper. These tools, which are not further described in this paper, facilitate the documentation, visualisation and navigation of ontologies, and enable them to be more readily understood by potential users such as academic researchers, publishers and librarians who, while expert in their own domains, may lack skill in ontology modelling and knowledge formalisation.

To inform our work, we have had informal discussions with publishers (including Oxford University Press, Public Library of Science, Royal Society of Chemistry and Società Editrice il Mulino), associated bodies (including ALPSP, CrossRef and Publishing Technologies) and members of academic communities (including the Digital Libraries, Document Engineering Linked Data and Semantic Web communities) to clarify their requirements.

We have attempted, as far as possible, to re-use existing models, ontologies and vocabularies of relevance to this domain, while developing our new modules to describe better those parts of the domain previously neglected.

The starting point for our work was version 1.6 of CiTO, the Citation Typing Ontology, described in [22]. This was both preliminary and incomplete, and yet contained within the single ontology both terms for handling bibliographic document descriptions and also properties to enable the characterisation of citations, as well as terms to permit recording of the number of citations to a given article, both within the citing paper and globally.

Our first decision, as we sought to develop ontological descriptions for this domain, was to modularise that ontology, in order to better develop each aspect of the original vision independently, and to keep each ontology module small and homogeneous enough to be readily understandable. In addressing this modularisation, we followed established methodologies for ontology modularization and development [19] [23], and took into account the following principles:

- The bibliographic description of documents and the characterisation of citations should be covered by two separate yet interoperable ontologies.
- To permit maximum reusability of each ontology module, logical constraints, for example domain and range constraints on properties, should be added only where they are strictly required.
- Where well-known and widely shared vocabularies covering parts of the domain we wanted to describe already existed, these should, where appropriate, be properly imported and re-used.

3 Related works

While wanting to re-use existing information models and vocabularies of relevance to scholarly publishing as far as possible, we had to come to terms with their limitations. In this section, we briefly introduce and comment upon these well-known vocabularies that we have considered and/or fully or partially employed within our own work.

Dublin Core. Born as consequence of a conference held in Dublin, Ohio, USA in 1995 that involved both technicians (librarians, publishers, archivists) and academics (researchers, software developers), the current versions of the Dublin Core (DC) Metadata Elements [9] and of the DC Metadata Terms [8] are the most widely used vocabularies for describing and cataloguing resources. While very useful for creating basic metadata that permit resource discovery, the main limitations of DC is a consequence of the *generic nature of its terms*.

PRISM. The *Publishing Requirements for Industry Standard Metadata* (*PRISM*) [14] is a specification defining a rich set of metadata terms for describing published works. The PRISM metadata terms are expressible both in XML, according to a specific DTD, and in RDF [11]. While PRISM has a much richer set of terms for describing bibliographic entities than DC, its main limitation is that it is a *flat structure*, lacking hierarchies. Developed primarily to describe the component parts of a bibliographic *record*, is not rich enough to describe the variety of bibliographic *entities* (e.g. research paper, book review, anthology) nor the relationships between them (e.g. that a research paper may be published as a journal article or as a conference paper; and that the same paper appearing as a journal article may have different formats - such as HTML and PDF - and different physical embodiments - as a computer file or as a ten-page printed paper). This requires one to use PRISM with another ontology to enable a complete description of bibliographic entities.

SKOS. Often, publishers need to classify the documents they publish according to discipline-specific thesauri or classification schemes, for example those belonging to economics[6] or computer science[7]. The *Simple Knowledge Organization System* (*SKOS*) [16] is an RDFS model to support the use of knowledge organization systems (KOS) such as thesauri, classification schemes, subject heading lists and taxonomies within the framework of the Semantic Web. The reception of this language has been particularly positive: a large number of well-known thesauri and classification systems have started to convert their respective specifications into SKOS documents[8]. This makes SKOS the *de facto*

4 Live OWL Documentation Environment (LODE): http://www.essepuntato.it/lode.

5 Graphical Framework For OWL Ontologies (Graffoo): http://www.essepuntato.it/graffoo. Note that Figures 1 and 2 of this article comply with Graffoo.

6 The Journal of Economic Literature Classification Scheme: http://www.aeaweb.org/jel/jel_class_system.php.

7 The Association for Computing Machinery (ACM) Computing Classification System 1998 http://www.acm.org/about/class/1998.

8 E.g. AGROVOC – http://aims.fao.org/website/AGROVOC-Thesaurus/sub; The Medical Subject Headings (MeSH) – http://www.ncbi.nlm.nih.gov/mesh; The Library of Congress Subject Headings (LCSH) – http://id.loc.gov/search/; and Nuovo

standard for encoding controlled vocabularies for the Semantic Web.

FOAF. The *Friend of a Friend Ontology* (FOAF)[9] is a simple and widely used vocabulary for describing people and their institutions [3], useful in this context for describing authors.

SWRC. The *Semantic Web for Research Communities* (*SWRC*) ontology [25] (labelled in German) enables the description of people, institutions, organisations and scholarly publications. It includes a few terms that describe bibliographic entities relevant to academia, such as journals, books and theses. However, since it has been developed primarily to describe academic research communities, its vocabulary covers only part of the publishing and bibliographic domain.

FRBR. The *Functional Requirements for Bibliographic Records* (*FRBR*) [13] is a general model, proposed by the International Federation of Library Association (IFLA), for describing documents and their evolution. It works for both physical and digital resources and has proven to be very flexible and powerful. FRBR describes all documents from four different and correlated points of view, as illustrated as follows by consideration of the book *Alice's Adventures in Wonderland* by Lewis Carroll:

- *Work*. A FRBR *Work* is a high-level abstract Platonic concept of the essence of a distinct intellectual or artistic creation, for example the ideas in Lewis Carroll's head concerning *Alice's Adventures in Wonderland*. A Work is realized through one or more Expressions.
- *Expression*. A FRBR *Expression* is the specific form that a *Work* takes each time its content is "realized" in physical or electronic form – thus the original text of *Alice's Adventures in Wonderland* and its Italian translation *Le Avventure di Alice nel Paese delle Meraviglie* refer to different expressions of the same work. An Expression is embodied in one or more Manifestations;
- *Manifestation*. A FRBR *Manifestation* of a work defines its particular physical or electronic embodiment, e.g. the particular *format* in which "Alice's Adventures in Wonderland" is stored, for instance as a printed object or in HTML. A Manifestation is exemplified by one or more Items.
- *Item*. A FRBR Item is a particular physical or electronic copy of *Alice's Adventures in Wonderland* that a person can own, e.g. the printed version of that book you have in your bookcase.

BIBO. BIBO [7] is an OWL Full ontology that allows one to write descriptions of documents for publication on the Semantic Web, *bibo:Document* being the core class of this model. It includes both DC and PRISM terms to cover common needs, uses FOAF to describe authors, and adds other classes and properties such as *bibo:AcademicArticle*, *bibo:Journal*, *bibo:Collection*, *bibo:Book*, *bibo:Chapter* and *bibo:Issue* to describe the publishing domain. BIBO is a good ontology that is widely used among the bibliographic community. A comparison of BIBO and FaBiO is given below, in Section 4.3.

SWAN Citations Ontology. Another model used to define bibliographic resources is the *Citations Ontology*[10] module included in the SWAN (Semantic Web Applications in Neuroscience) Ontologies, version 1.2 [6]. In this ontology, compliant with OWL 2 DL, the main class *Citation*[11] is used as super-class, of which everything else (e.g. *JournalArticle*, *WebArticle* and *Book*) is sub-class.

4 Representing bibliographic information using FaBiO

The need for ontologies that are sufficiently expressive for describing documents has been presented in Section 1.

The vocabularies described above are either poor in concepts or are 'flat', preventing their use for accurately describing the complexity of publishing reality. We will illustrate this by considering the representation of a typical bibliographic reference using first BIBO and then FRBR. We will then show how this information can be more accurately described using FaBiO. Consider the following typical bibliographic reference:

> Yves Marcoux, Élias Rizkallah (2009). Intertextual semantics: A semantics for information design. *Journal of the American Society for Information Science and Technology*, 60 (9): 1895-1906. John Wiley & Sons, Inc. DOI: 10.1002/asi.21134. First published online (PDF and HTML) 21 August 2009.

From the previous description we can extract the following information:

1. The document is an academic research article – deducible from the journal in which it is published.
2. Yves Marcoux and Élias Rizkallah are the authors of the article.
3. The article was published in 2009.
4. The article is entitled "Intertextual semantics: A semantics for information design".
5. It was published in the 9th Issue of the 60th volume of the *Journal of the American Society for Information Science and Technology*.
6. The DOI of the article is "10.1002/asi.21134".
7. The article was first published online on the 21st of August, 2009, in two particular formats (PDF and HTML).
8. The page range of the article within the printed version is from starting page 1895 to ending page 1906.
9. The publisher of the journal is John Wiley & Sons, Inc.

4.1 Bibliographic reference metadata encoding using BIBO

BIBO [7] is the first OWL ontology specifically designed to address the domain under discussion, and it expands the DC Terms vocabulary with terms specific for bibliographic metadata, particularly relating to legal documents, and also for various types of events. It includes PRISM and FOAF terms.

Soggettario of the National Central Library in Florence – http://thes.bncf.firenze.sbn.it/.

9 FOAF: http://www.foaf-project.org/; http://xmlns.com/foaf/spec/.

10 SWAN Citations Ontology Module: http://swan.mindinformatics.org/spec/1.2/citations.html.

11 Note that in this ontology the word "Citation" is used to denote the **object** of a citation, i.e. a cited document, rather than the performative act of making a citation, which is the usage employed both in CiTO and within this article.

In the following RDF encoding example, the information given in the bibliographic reference cited above is encoded using BIBO[12]:

```
:intertextual-semantics a bibo:AcademicArticle
  ; bibo:authorList ( :marcoux :rizkallah )
  ; dcterms:title "Intertextual semantics: A
semantics for information design"
  ; dcterms:issued "2009"^^xsd:gYear
  ; bibo:doi "10.1002/asi.21134"
  ; bibo:pageStart "1895"
  ; bibo:pageEnd "1906"
  ; dcterms:hasFormat :html , :pdf
  ; dcterms:isPartOf [ a bibo:Issue
    ; bibo:issue "9"
    ; bibo:volume "60"
    ; dcterms:isPartOf [ a bibo:Journal
      ; dcterms:title "Journal of the American
Society for Information Science and
Technology"
      ; dcterms:publisher :wiley-and-sons ]] .

:html a bibo:AcademicArticle
  ; dcterms:format type:text/html
  ; dcterms:issued "21-08-2009"^^xsd:date .

:pdf a bibo:AcademicArticle
  ; dcterms:format type:application/pdf
  ; dcterms:issued "21-08-2009"^^xsd:date .

:marcoux a foaf:Person
  ; foaf:givenName "Yves"
  ; foaf:familyName "Marcoux" .

:rizkallah a foaf:Person
  ; foaf:givenName "Élias"
  ; foaf:familyName "Rizkallah" .

:wiley-and-sons a foaf:Organization
  ; foaf:name "John Wiley & Sons, Inc." .
```

There are some obscure points that emerge from the preceding formalisation:

- Some of the statements are *too generic*. For example, the property *dcterms:issued* that is used to represent the various dates associated with the publication of this article, is itself employed in conjunction with two different date formats, i.e. the publication year `"21-08-2009"^^xsd:date` and the date the article was first published online `"2009"^^xsd:gYear`. While this design choice is reasonable, it lacks precision.
- BIBO specifies that the property for listing authors (*bibo:authorList*) must have, as its range, either an *rdf:List* or an *rdf:Seq*, which implicitly results in having an OWL full ontology. Of course this is a perfectly reasonable and valid modelling choice. However, it would prevent potential use of DL-oriented reasoners over data encoded in this way, since OWL 2 DL does not support these RDF classes[13].
- Like PRISM, BIBO can record a volume number through the data property *bibo:volume*, but, although it has the classes *bibo:AcademicArticle*, *bibo:Issue* and *bibo:Journal*, BIBO lacks the concept of "Volume" as a distinct class among other bibliographic classes that have a hierarchical partitive relationship to one another (i.e. Journal Article > Issue > Volume > Journal).
- The relations between the various formats of the article in the example are not clear. For example, the manner in which the resource "*:intertextual-semantics*" relates to the resource "*:pdf*" is not specified. Does the latter represent the content of the former in a particular format, or there is something more?
- Additionally, there is no precise way to indicate that the page numbers refer only to the printed article, not to the HTML version.
- Furthermore, because it lacks the layered structure of FRBR, BIBO does not have the flexibility to distinguish between concepts at these various levels, for example that an academic paper could have various possible expressions as a journal article, a conference paper or a book chapter. The class *bibo:AcademicArticle* is in fact a conflation of the concepts "academic paper" and "journal article".

4.2 Bibliographic reference metadata encoding using FRBR

It is possible to handle some of the issues raised above by adopting the more structured FRBR model [13], as expressed in the FRBR Core ontology[14], together with DC terms for textual statements (i.e. those statements having a literal string as their object). This is illustrated in the following example:

```
:intertextual-semantics a frbr:Work
  ; frbr:creator :marcoux , :rizkallah
  ; frbr:realization :content .

:content a frbr:Expression
  ; dcterms:title "Intertextual semantics: A
semantics for information design"
  ; dcterms:issued "2009"^^xsd:gYear
  ; dcterms:identifier "doi:10.1002/asi.21134"
  ; frbr:embodiment :printed , :html , :pdf
  ; frbr:partOf [ a frbr:Expression
    ; dcterms:identifier "9"
    ; dcterms:description "Issue"
    ; frbr:embodiment :printed-issue
    ; frbr:partOf [ a frbr:Expression
      ; dcterms:identifier "60"
      ; dcterms:description "Volume"
      ; frbr:partOf [ a frbr:Expression
        ; dcterms:title "Journal of the
American Society for Information Science and
Technology" ] ] ] .

:printed-issue a frbr:Manifestation
  ; frbr:producer :wiley-and-sons
  ; dcterms:issued "09-2009"^^xsd:gYearMonth
  ; frbr:part :printed .

:printed a frbr:Manifestation
  ; frbr:producer :wiley-and-sons
  ; dcterms:issued "09-2009"^^xsd:gYearMonth
  ; dcterms:extent [ a dcterms:SizeOrDuration
      ; dcterms:description "1895-1906"  ] .

:html a frbr:Manifestation
  ; frbr:producer :wiley-and-sons
  ; dcterms:format type:text/html
  ; dcterms:issued "21-08-2009"^^xsd:date .
```

12 This and all the following RDF excerpts are written in Turtle [2].

13 For a longer and clearer justification of why RDF collections and containers are not usable and interpreted correctly by OWL 2 DL, please consult http://hcklab.blogspot.com/2008/12/moving-towards-swan-collections.html.

14 The FRBR Core in RDFS and the related implementation in OWL 2 DL are available respectively at http://vocab.org/frbr/core and http://purl.org/spar/frbr. Note that in the latter, *frbr:part* and *frbr:partOf* are both defined as transitive object properties.

```
:pdf a frbr:Manifestation
  ; frbr:producer :wiley-and-sons
  ; dcterms:format type:application/pdf
  ; dcterms:issued "21-08-2009"^^xsd:date .

:marcoux a frbr:Person
  ; dcterms:description "Yves Marcoux" .

:rizkallah a frbr:Person
  ; dcterms:description "Élias Rizkallah" .

:wiley-and-sons a frbr:CorporateBody
  ; dcterms:description
    "John Wiley & Sons, Inc." .
```

Although it is possible to use FRBR in this manner to give a structured and unambiguous description of all the bibliographic entities, this example makes clear the severe limitations of FRBR, due to the lack of specific terms in the FRBR Core ontology to permit publications to be described in normal everyday language, e.g. "Journal" or "Page".

4.3 Bibliographic reference metadata encoding using FaBiO

FaBiO, the *FRBR-aligned Bibliographic Ontology* (current version 1.6), was developed precisely to address the issues revealed by the previous examples, while re-using the previous fundamental work in this domain (so as not to re-invent the wheel). In particular, DC Terms, PRISM, FRBR and SKOS terms are all included in our model.

Considering again the previous bibliographic reference example, a possible FaBiO formalisation is:

```
:intertextual-semantics a fabio:ResearchPaper
  ; dcterms:creator :marcoux , :rizkallah
  ; frbr:realization :version-of-record .

:version-of-record a fabio:JournalArticle
  ; dcterms:title "Intertextual semantics: A
semantics for information design"
  ; fabio:hasPublicationYear "2009"^^xsd:gYear
  ; prism:doi "10.1002/asi.21134"
  ; frbr:embodiment :printed , :html ,:pdf
  ; frbr:partOf [ a fabio:JournalIssue
    ; prism:issueIdentifier "9"
    ; frbr:embodiment :printed-issue
    ; frbr:partOf [ a fabio:JournalVolume
      ; prism:volume "60"
      frbr:partOf [ a fabio:Journal
        ; dcterms:title "Journal of the
American Society for Information Science and
Technology" ] ] ] .

:printed a fabio:PrintObject
  ; prism:startingPage "1895"
  ; prism:endingPage "1906"
  ; frbr:part of :printed-issue .

:printed-issue a fabio:Paperback
  ; dcterms:publisher :wiley-and-sons
  ; prism:publicationDate
    "09-2009"^^xsd:gYearMonth .



:html a fabio:WebPage
  ; dcterms:publisher :wiley-and-sons
  ; dcterms:format type:text/html
  ; prism:publicationDate
    "21-08-2009"^^xsd:date .

:pdf a fabio:DigitalManifestation
  ; dcterms:publisher :wiley-and-sons
  ; dcterms:format type:application/pdf
  ; prism:publicationDate
    "21-08-2009"^^xsd:date .

:marcoux a foaf:Person
  ; foaf:givenName "Yves"
  ; foaf:familyName "Marcoux" .

:rizkallah a foaf:Person
  ; foaf:givenName "Élias"
  ; foaf:familyName "Rizkallah" .

:wiley-and-sons a foaf:Organization
  ; foaf:name "John Wiley & Sons, Inc." .
```

(Note that *fabio:WebPage* is subclass of *fabio:DigitalManifestation,* that *fabio:Paperback* is a subclass of *fabio:PrintObject*, and that *fabio:PrintObject* is subclass of *fabio:AnalogManifestation*).

With FaBiO, it thus becomes possible:

- To write semantic descriptions of a wide variety of bibliographic objects, including research articles, journal articles, journal issues and journal volumes, using terms that closely resemble the language used in everyday speech by academics and publishers;
- To employ FRBR categories to define clear separations between each part of the publishing process, involving different people (authors, publishers, readers) depending on which aspect of the bibliographic entity we are considering: the high-level conceptualisation of the research paper, the version of record of that paper forming a journal article, the publication of that article in various formats, and the individual physical or electronic exemplars of the published article that people may read and own.
- To include with ease elements from other vocabularies for describing particular entities involved in a publishing process that are not specified by FaBiO itself, such as those from FOAF for persons and organizations.

The most closely similar ontology to FaBiO in terms of its terms and intended usage is *BIBO*, the *Bibliographic Ontology*, which has already been introduced in Section 3 and in Section 4.1. The principal differences between BIBO and FaBiO are:

1. that FaBiO is structured according to the FRBR model, providing the greater expressivity required for unambiguously describing the various essences of a bibliographic object, as discussed above[15]; and
2. that it has a more comprehensive collection of classes and properties than does BIBO, permitting more precise descriptions of bibliographic entities.

This semantic richness of FaBiO is enhanced when it is used together with other SPAR ontologies, such as the *Citation Typing Ontology* (CiTO), described in Section 5, the

15 Note that understanding when an authorial process produces a new Work, or just another Expression of a previous Work, strictly depends on the particular scenario under consideration. FRBR does not give any strict guideline to follow in these cases – and neither does FaBiO. Thus, for instance, the relation between a conference paper and a subsequent journal article can differ case by case. It can happen that the journal article is just another Expression of a common Work (e.g. if we *revise* the original text of the conference paper by correcting just few typos or misworded paragraphs), or alternatively an entirely new Work (e.g. when we adapt and extend the conference paper in a significant way to make it suitable for journal publication).

Bibliographic Reference Ontology (BiRO)[16] for describing bibliographic references, reference lists, bibliographic records, etc., the *Document Components Ontology* (DoCO)[17] for describing the structural and rhetorical components of documents (e.g. paragraph, Results), the *Publishing Roles Ontology* (PRO)[18] for defining roles held by agents involved in a publication process (e.g. author, publisher, editor, reviewer), and the *Publishing Status Ontology* (PSO)[19] for defining the possible status of a bibliographic object (e.g. submitted, under review, accepted for publication) [17].

BIBO might have been considered a good starting point for the development of FaBiO. However, the fact that BIBO collapses everything describing a bibliographic entity within the classes *bibo:Document* or *bibo:Collection* makes it impossible to align BIBO classes with FRBR endeavours (Works, Expressions, Manifestations and Items, that are all disjointed classes) while keeping the entire ontology logically consistent.

It should perhaps be mentioned here that we deliberately excluded a class "*fabio:Document*" from FaBiO, since the term "Document" has so many different possible definitions as to make its inclusion within the FRBR hierarchy problematic. We have not found this exclusion to pose any problems when using FaBiO to describe bibliographic entities.

It is thus for users to choose which of these two ontologies, BIBO or FaBiO, best suits their purposes. Because of their similarities, we have defined a SKOS alignment between BIBO and FaBiO, in order to facilitate metadata conversion between the two models. That alignment document, entitled *BIBO2SPAR*, is freely available online[20].

Other advantages of FaBiO are outlined below.

4.4 New capabilities when using FaBiO

4.4.1 Using external models

As already mentioned, FaBiO was developed with the minimum of restrictions to its classes and to the domains and ranges of its properties. This flexibility has the great advantage of allowing FaBiO to be used together with other ontologies. We have already seen how FOAF can be used to describe agents. Another common requirement is that of specifying the order of components in a list, for example authors in an author list or references in a reference list. Unlike the use of *bibo:authorList*, which breaks OWL 2 DL compliance as explained above, this can be achieved in a manner that is compliant with the decidable and computable OWL 2 DL by combining FaBiO with the Collections Ontology (CO)[21], an OWL 2 DL ontology specifically designed for defining orders among items, in the following way:

```
:intertextual-semantics a fabio:ResearchPaper
  ; dcterms:creator :listOfAuthors .

:listOfAuthors a co:List
  ; co:firstItem [co:itemContent :marcoux
  ; co:nextItem [co:itemContent :rizkallah ]].
```

In this way we can still keep the model in OWL 2 DL. Additionally, because the ranges of *dcterms:creator* and other properties within FaBiO have intentionally been left unspecified, FaBiO guarantees a level of interoperation with other models without incurring in any undesirable collateral effects, such as ontology inconsistencies or the generation of undesired inferences.

For instance, OAI-ORE [12] is excellent well-known vocabulary that can be combined with FaBiO to describe aggregations of bibliographic entities, contributors and the like, since FaBiO and OAI-ORE are both very flexible and reusable. However, we did not want to restrict users to using only OAI-ORE, instead of other models (e.g. CO), to define these kinds of scenarios[22].

4.4.2 Extending FRBR within FaBiO

One of the explicit requests from publishers and end-users was to be able to create shortcuts between FRBR endeavours (work, expression, manifestation, item) that were not part of the original FRBR model. Let us introduce an example to illustrate these needs, by slightly changing the bibliographic reference we introduced previously:

> Yves Marcoux, Élias Rizkallah (2009). Intertextual semantics: A semantics for information design. http://onlinelibrary.wiley.com/doi/10.1002/asi.21134/full.

In this reference, we have just a FRBR work – the paper by Marcoux and Rizkallah – and the URL for a specific FRBR item that portrays that work – the HTML version of the paper on the publisher's web site. If we wished to link these concepts using the FRBR OWL ontology terms we have employed so far, we would be obliged to specify each intermediate FRBR endeavour, namely the expression and manifestation of that paper, even if we were not interested in doing so:

```
:intertextual-semantics a frbr:Work
  ; frbr:creator :marcoux , :rizkallah
  ; frbr:realization [ a frbr:Expression
    ; dcterms:title "Intertextual semantics: A
semantics for information design"
    ; frbr:embodiment [ a frbr:Manifestation
      ; frbr:exemplar
        wiley:10.1002/asi.21134/full ] ] .
```

In order to avoid this kind of verbosity, it is possible to use the new FaBiO properties shown in Figure 1 to link a work directly to its manifestations (*fabio:hasManifestation*) or to its items (*fabio:hasPortrayal*), and to link an expression directly to its items (*fabio:hasRepresentation*).

[Figure 1: PLEASE INSERT HERE]

These added properties allows us to be less verbose:

```
:intertextual-semantics a frbr:Work
  ; frbr:creator :marcoux , :rizkallah
  ; frbr:realization [ a frbr:Expression
    ; dcterms:title "Intertextual semantics: A
semantics for information design"
    ; fabio:hasRepresentation
      wiley:10.1002/asi.21134/full .
```

16 The Bibliographic Reference Ontology (BiRO): http://purl.org/spar/biro.

17 The Document Components Ontology (DoCO): http://purl.org/spar/doco.

18 The Publishing Roles Ontology (PRO): http://purl.org/spar/pro.

19 The Publishing Status Ontology (PSO): http://purl.org/spar/pso.

20 BIBO2SPAR mapping document: http://sempublishing.svn.sourceforge.net/svnroot/sempublishing/SPAR/BIBO2SPAR.ttl.

21 CO, the Collections Ontology: http://purl.org/co.

22 Examples of how to use FaBiO and OAI-ORE in combination lie beyond the scope of this paper.

4.4.3 Categorising bibliographic resources with SKOS

One of the most important needs for a publisher is to categorise each bibliographic entity it produces by adding free-text keywords and/or specific terms structured according to recognised classification systems and/or thesauri. While the definition of keywords is possible in FaBiO using the PRISM property *prism:keyword*, terms from thesauri, structured vocabularies and classification systems are best described using SKOS [16]. To facilitate this, FaBiO extends some classes and properties of SKOS as shown in Figure 2.

[Figure 2: PLEASE INSERT HERE]

As shown, any FRBR endeavour can be associated (*fabio:hasSubjectTerm*) with one or more descriptive terms (*fabio:SubjectTerm*, a sub-class of *skos:Concept*) found in a specific dictionary (*fabio:TermDictionary*, a sub-class of *skos:ConceptScheme*) that is relevant to (*fabio:hasDiscipline*) one or more particular disciplines (*fabio:SubjectDiscipline*, also a sub-class of *skos:Concept*) describing a field of knowledge or human activity such as Computer Science, Biology, Economics, Cookery or Swimming. At the same time, the subject disciplines themselves can be grouped into a *fabio:DisciplineDictionary*.

Our selected example can thus be enriched as follows:

```
:intertextual-semantics a fabio:ResearchPaper
  ; fabio:hasSubjectTerm acm:markup-languages
    , acm:semantics
  ; prism:keywords "semantics of markup"
    , "semiotic application" , "xml" .

<http://www.acm.org/class/1998>
  a fabio:TermDictionary
  ; skos:prefLabel "The 1998 ACM Computing
Classification System"
  ; fabio:hasDiscipline
    dbpedia:Computer_science .

acm:markup-languages a fabio:SubjectTerm
  ; skos:prefLabel "Markup languages"
  ; skos:inScheme
    <http://www.acm.org/class/1998>
  ; skos:broader acm:document-preparation .

acm:semantics a fabio:SubjectTerm
  ; skos:prefLabel "Semantics"
  ; skos:inScheme
    <http://www.acm.org/class/1998>
  ; skos:broader
    acm:formal-definitions-and-theory .
```

5 Characterising citations with CiTO

Bibliographic citation, i.e., the act of referring from a citing entity to the cited one, is one of the most important activities of an author in the production of any bibliographic work, since the acknowledgement of sources that this activity represents stands at the very core of the scholarly enterprise. The network of citations created by combining citation information from many academic articles and books is a source of rich information for scholars, and can be used by publisher to create new and interesting ways of browsing their data, as well as for calculating metrics reflecting the relative importance of a journal (e.g. the *impact factor*) or an author (e.g. the *h-index*).

The reasons that an author cites other publications are manifold, as pointed out by Teufel *et al.* [26] and Jörg [15]. Usually, it is because the author has gained assistance of some sort, perhaps in the form of background information, ideas, methods or data, from the cited previously published works, and wishes to acknowledge this. More rarely, citations may be made because the citing works review, critique or refute previous works. Most citations are direct and explicit (as in the reference list of a journal article). However, they can also be indirect (for example, by means of a citation to a more recent paper by the same research group on the same topic, e.g. as a means of citing a method first described in an earlier preliminary paper), or implicit (as in artistic quotations or parodies, or *in extremis* in cases of plagiarism).

Classical scholarship has well-developed methods for citing individual sections, paragraphs or verses of referenced works. In contrast, in modern scientific practice a citation is made to the previously published paper as a whole, with little or no indication given as to why that paper has been cited or what portions of it are relevant to the discussion at hand, except what the reader can glean from the citation context.

Of course, previously developed models for describing bibliographic objects allow for the existence of citations among bibliographic entities to be recorded. For instance, taking again the previous example about the article "Intertextual semantics: A semantics for information design", using BIBO [7] we can declare citations as follows:

```
# :version-of-record is the published
# Expression of :intertextual-semantics
:version-of-record bibo:cites
  # The following resources are Expressions
  # ('vor' stands for 'version of record')
  :towards-a-semantics-vor
  , :meaning-and-interpretation-vor
  , :design-everyday-things-vor
  , :exploring-intertextual-semantics-vor …
```

Alternatively, we can use the Discourse Relationships Module[23] in SWAN v1.2 [6] in the same way:

```
: version-of-record
  discourse-relationships:cites
    :towards-a-semantics-vor
    , :meaning-and-interpretation-vor
    , :design-everyday-things-vor
    , :exploring-intertextual-semantics-vor …
```

However, while the *cites* properties in these two ontologies, as well as the more generic property *dcterms:relation* in DC Terms, permit the bald *existence* of a citation to be recorded, they do not permit the author to augment the citation with any specific factual and/or rhetorical meanings that would describe the *reasons* the author had in mind when creating citations to some particular documents rather than to others.

CiTO, the *Citation Typing Ontology*[2], seeks to improve upon this situation by making it possible for authors (or others) to capture the intent when citing a particular publication, permitting them to create metadata describing their citations that is quite distinct from metadata describing the cited works themselves. CiTO thus permits the motivations of an author when referring to another document to be captured. For instance, the previous example may be rewritten using CiTO as follows:

```
: version-of-record
  cito:disputes :towards-a-semantics-vor
  ; cito:citesAsRelated
    :meaning-and-interpretation-vor
  ; cito:agreesWith
    :design-everyday-things-vor
  ; cito:extends
    :exploring-intertextual-semantics-vor …
```

The current version of CiTO, version 2.3, contains and extends the citation-specific object properties that were

23 The Discourse Relationships Ontology: http://swan.mindinformatics.org/spec/1.2/discourserelationships.html.

originally contained in CiTO version 1.6 [22], to the exclusion of all other original classes and properties within CiTO v1.6, which, as part of the modularization we have undertaken, have been moved into FaBiO or into two of the other SPAR ontologies, the *Citation Counting and Context Characterization Ontology* (C4O)[24] and the *Publishing Status Ontology* (PSO)[19].

CiTO now contains just two main object properties, *cito:cites* and its inverse *cito:isCitedBy*, each of which has thirty-two sub-properties, plus two additional generic object properties that may be used to make statements relating two entities that do not constitute formal citation acts: *cito:shareAuthorsWith* and *cito:likes*, the latter permitting social media 'likes' statements to be encoded in RDF.

As shown in Figure 3, these properties (and, consequently, their inverses) may be classified as rhetorical[25] and/or factual, with the rhetorical properties being grouped in three sets depending on their connotation: *positive*, *informative* (or *neutral*) or *negative*[26].

[Figure 3: PLEASE INSERT HERE]

When developing CiTO from CiTO v1.6, we intentionally removed the domain and range constraints from its object properties, so that this ontology could be easily integrated with other models. Obviously, it can be successfully used in conjunction with FaBiO, so that descriptions of a bibliographic entity and its citations can be mixed within a single RDF graph:

```
:version-of-record a fabio:JournalArticle
  ; dcterms:title "Intertextual semantics: A
semantics for information design"
  ; cito:disputes :towards-a-semantics-vor …

:towards-a-semantics-vor fabio:ConferencePaper
  ; :frbr:realizationOf [
    a fabio:ResearchPaper
    ; dcterms:creator :renear
      , :dubin , :sperberg-mcqueen ]
  ; dcterms:title
    "Towards a semantics for XML markup"
  ; cito:isDisputedBy : version-of-record …
```

In addition, we have recently included in CiTO additional properties (i.e. *cito:compiles* and *cito:isCompiledBy*) for making specific assertions about datasets, computer programs and other similar digital objects, and we have also developed the Argument Model Ontology[27], an ontology that enables descriptions of argumentation in scholarly works according to the Toulmin model of argument [27].

6 Community uptake of FaBiO and CiTO

FaBiO and CiTO are now being used or are being considered for adoption in a variety of academic and publishing environments, as described below. The adoption of these models by different communities can be ascribed, at least in part, to the minimization of the constraints applied to the ontological entities, so that the ontologies can be applied in a wide variety of situations.

Linked Open Vocabularies Dataset. The Linked Open Vocabularies Dataset (LOV)[28], that describes the growing ecosystem of linked open vocabularies (RDFS or OWL ontologies) used in the Linked Data Cloud, gives the SPAR Ontologies a special node within the Library vocabulary space[29].

SWAN ontology. The most recent version (v2.0) of the *SWAN ontology ecosystem* [6] has recently been harmonised to include FaBiO and work seamlessly with CiTO [5]. This harmonisation was undertaken collaboratively by the authors, while developing version 1.6 of CiTO [22] into CiTO v2.0 and FaBiO v1.0, and by Paolo Ciccarese and Tim Clark (Harvard University), authors of the SWAN Ontologies. The resulting combined CiTO/FaBiO + SWAN model is specified in OWL 2 DL, is fully modular, and inherently supports agent-based searching and mash-ups.

CiteULike. Egon Willighagen has pioneered the use of CiTO[30] to characterize bibliographic citations within *CiteULike*[31], the free service for managing and discovering scholarly references. A user can add a CiTO relationship between articles via the CiteULike interface, provided that both the citing and the cited articles are in the user's library.

WordPress. Martin Fenner describes[32] a plug-in for *WordPress* called *Link-to-Link*[33], that makes it easy to add citation typing to references within a blog post, using a sub-set of the most commonly used CiTO relationships presented in a convenient drop-down menu.

Linked Education. The open platform *Linked Education*[34], which aims at sharing and promoting the use of Linked Data for educational purposes, recently added CiTO to its listing[35] of RDF schemas and vocabularies suitable for use in educational contexts, for example to describe educational resources.

Virtual Observatory. Accomazzi and Dave [1] report the adoption of the FaBiO and CiTO ontologies as part of their semantic knowledge base, allowing easier integration and linking of the body of heterogeneous astronomical resources into what they term a Virtual Observatory.

Open Citations Corpus. The Open Citations Corpus[36] is a database of approx. 6.3 million biomedical literature citations, harvested from the reference lists of all open access articles in PubMed Central. These contain references to approx. 3.4 million papers, which represent ~20% of all PubMed-listed papers published between 1950 and 2010, including all the most highly cited papers in every biomedical field. CiTO, FaBiO and other SPAR ontologies have been

24 C4O, the Citation Counting and Context Characterization Ontology: http://purl.org/spar/c4o.

25 We use the word "rhetorical" to refer to the particular (subjective) role and function that a portion of text and/or a bibliographic entity may have within a scholarly publication, as described, for example, by [10], [20] and [29].

26 Although we cluster the CiTO properties in Figure 3, their formal hierarchy is left intentionally vague. Since the exact meanings of terms may be somewhat overlapping, we think this lack of rigid structure increase the reusability of CiTO properties among different contexts.

27 The Argument Model Ontology: http://www.essepuntato.it/2011/02/argumentmodel.

28 LOV, The Linked Open Vocabularies Dataset: http://labs.mondeca.com/dataset/lov/.

29 SPAR in LOV: http://labs.mondeca.com/dataset/lov/details/vocabularySpace_SPAR.html.

30 http://chem-bla-ics.blogspot.com/2010/10/citeulike-cito-use-case-1-wordles.html.

31 CiteULike: http://www.citeulike.org/.

32 http://blogs.plos.org/mfenner/2011/02/14/how-to-use-citation-typing-ontology-cito-in-your-blog-posts/.

33 http://wordpress.org/extend/plugins/link-to-link/.

34 http://linkededucation.org/.

35 http://linkededucation.wordpress.com/data-models/schemas/.

36 The Open Citations Corpus: http://opencitations.net/.

used to encode information about bibliographic records in RDF, as explained in the Open Citations Blog[37].

Web Tracks. WebTracks[38] is an open source project funded by the JISC Managing Research Data Programme[39] that is developing a peer-to-peer protocol to enable web-scale link tracking. WebTracks focuses on actually making these connections, particularly between research datasets and related publications. WebTracks creates semantically annotated links between data resources using CiTO, yielding a graph of citation and provenance to enable web-scaled data management by exposing links between related objects.

Società Editrice il Mulino. The active collaboration between the Italian scholarly publishing house *Società Editrice il Mulino*[40] and the Department of Computer Science of the University of Bologna has led to the recent prototyping of an application called *Folksauro* (the name comes from the concatenation of the words *folksonomy* and *thesaurus*). Using FaBiO and DoCO as its main ontologies, and one or more discipline-specific thesauri developed in SKOS, Folksauro allows a user to associate terms from the thesauri and/or free-text keywords with the whole document, and/or with its sub-parts (chapters, sections, paragraphs, etc.), by means of an intuitive interface that hides the complexity of models and languages used.

Utopia. Utopia Documents[41] is a novel "smart PDF reader" that semantically integrates visualization and data-analysis tools with published research articles in PDF format. It uses CiTO to record bibliographic citations, and has a mechanism that deconstructs a PDF document into its constituent parts, which are then annotated using DoCO, the Document Components Ontology[17]. It also pulls bibliographic citation data live from the Open Citations Corpus and uses them to display the citation network for the paper manifested by the PDF, or for any of the articles referenced in that paper's reference list.

CrowdoMeter. CrowdoMeter[42] is a project by Euan Adie and Martin Fenner. It displays tweets about scientific articles and allows one to link those with the original articles using a subset of CiTO properties.

Data.open.ac.uk. The Open University has recently developed a platform[43] that allows the extraction, interlinking and publication of data concerning the institutional repositories of the Open University, with the intention of making them available for reuse and querying. All the citations between scholarly papers present in the dataset are described using CiTO.

7 Conclusions

In this article we have presented FaBiO, the *FRBR-aligned Bibliographic Ontology* (current version 1.6), and CiTO, the *Citation Typing Ontology* (current version 2.3). FaBiO and CiTO are two new OWL 2 DL ontologies for describing bibliographic resources and bibliographic citations on the Semantic Web. We have introduced those two models step by step, in order to emphasise their features for describing bibliographic objects and to stress their advantages relative to other pre-existing models. Finally, we have reviewed related work, and have described the community uptake of FaBiO, CiTO and other SPAR ontologies.

We plan to extend our contribution in this field from different perspectives. First of all, we are working to persuade publishers to adopt FaBiO and CiTO for recording and exposing their bibliographic and reference metadata. At the same time, we plan to extend CiTO to include some additional citation categories, particularly *cito:contrastsWith*, bearing in mind the citation annotation scheme developed by Teufel *et al*. in [26], which, while in broad agreement with our own, has a different viewpoint, uses a different vocabulary and introduces some new concepts. In addition, our intention is now to develop a community around the SPAR ontologies, to maintain their development and expand them to cover other scenarios within the publishing domain, particularly the publication of datasets. We welcome contacts from those who would like to participate in this work.

37 The Open Citations Blog: http://opencitations.wordpress.com/.

38 WebTracks: http://webtracks.jiscinvolve.org/wp/about/.

39 JISC Managing Research Data Programme: http://www.jisc.ac.uk/whatwedo/programmes/mrd.aspx.

40 Società Editrice il Mulino: http://www.mulino.it.

41 Utopia Documents: http://getutopia.com.

42 Crowdometer: http://crowdometer.org.

43 Open Linked Data from The Open University: http://data.open.ac.uk.

Acknowledgements

We thank Paolo Ciccarese and Tim Clark of Harvard University for their support and valuable conceptual and technical suggestions given while we were developing FaBiO and CiTO from CiTO v1.6 and were collaborating with them to harmonize these with SWAN. We are also grateful to our local colleagues, particularly Jun Zhao, Graham Klyne and Fabio Vitali, for their warm support, constructive criticisms and help given throughout these developments. Aspects of this work have been supported by the JISC (Joint Informations Systems Committee) through a grant to DS to support the Open Citations Project. This work would not have come about but for the enlightened policy of the Department of Computer Science of the University of Bologna of sending doctoral students on extended internships in foreign universities, from which both authors have greatly benefited.

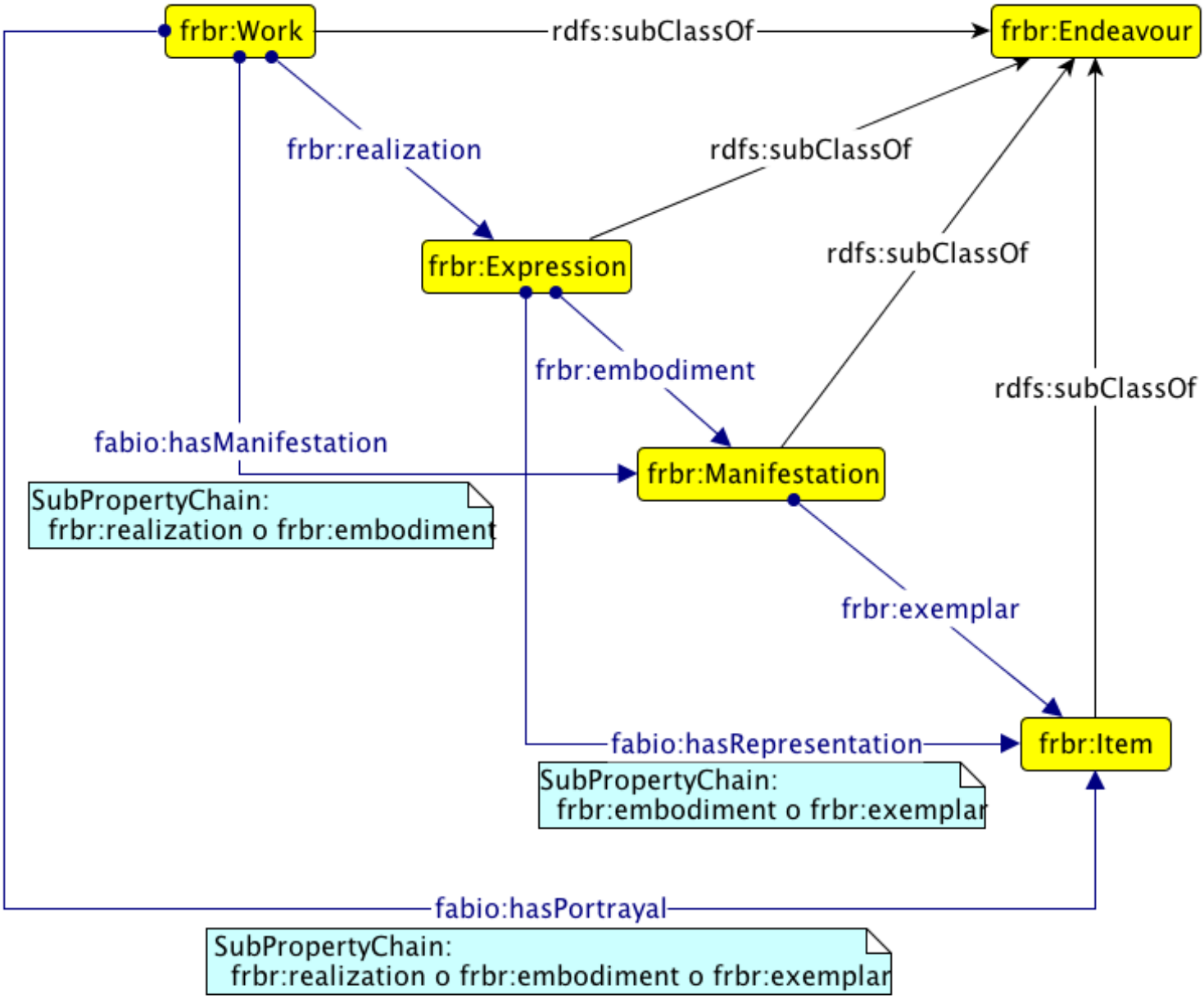
References

1. Accomazzi, A., Dave, R. (2011). Semantic Interlinking of Resources in the Virtual Observatory Era. ArXiv:1103.5958, http://arxiv.org/abs/1103.5958 (last visited 25th July 2012).
2. Becket, D., Berners-Lee, T. (2011). Turtle - Terse RDF Triple Language. W3C Team Submission, 28 March 2011. World Wide Web Consortium. http://www.w3.org/TeamSubmission/turtle/ (last visited 25th July 2012).
3. Brickley, D., Miller, L. (2010). FOAF Vocabulary Specification 0.98. Namespace Document, 9 August 2010 - Marco Polo Edition. http://xmlns.com/foaf/spec/ (last visited 25th July 2012).
4. Carroll, J., Klyne, G. (2004). Resource Description Framework (RDF): Concepts and Abstract Syntax. W3C Recommendation, 10 February 2004. World Wide Web Consortium. http://www.w3.org/TR/rdf-concepts/ (last visited 25th July 2012).
5. Ciccarese, P., Shotton, D., Peroni, S., Clark, T. (2011). CiTO + SWAN: The Web Semantics of Bibliographic Records, Citations, Evidence and Discourse Relationships. *Semantic Web Journal*. Special Issue: Semantic Web – Interoperability,

Usability, Applicability. Available at http://www.semantic-web-journal.net/content/cito-swan-web-semantics-bibliographic-records-citations-evidence-and-discourse-relationships (last visited 25th July 2012).

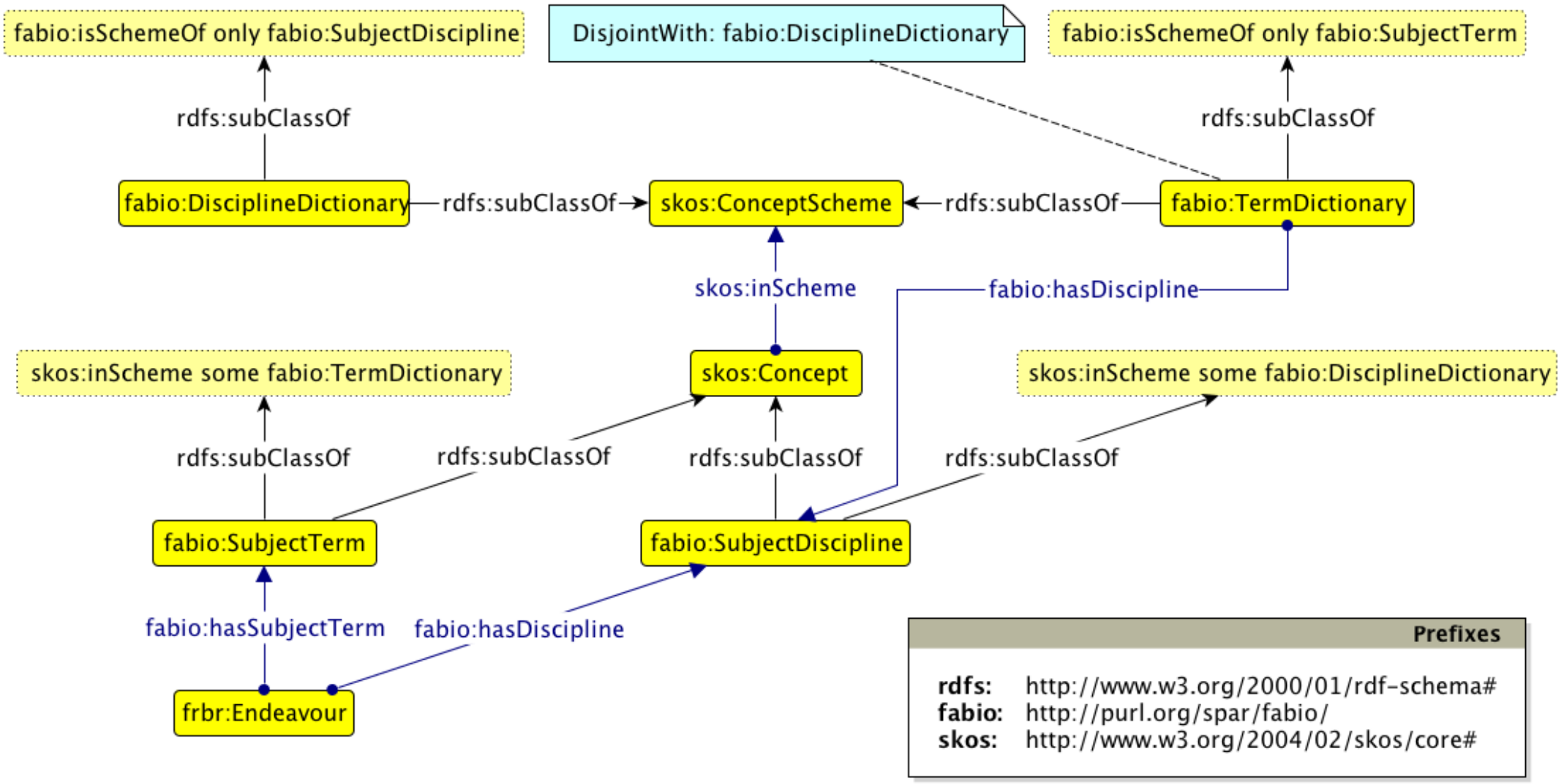
6. Ciccarese, P., Wu, E., Kinoshita, J., Wong, G., Ocana, M., Ruttenberg, A., Clark, T. (2008). The SWAN Biomedical Discourse Ontology. *Journal of Biomedical Informatics*, 41 (5), 739-751. doi:10.1016/j.jbi.2008.04.010.
7. D'Arcus, B., Giasson, F. (2009). Bibliographic Ontology Specification. Specification Document, 4 November 2009. http://bibliontology.com/specification (last visited 25th July 2012).
8. Dublin Core Metadata Initiative (2010). DCMI Metadata Terms. DCMI Recommendation. http://dublincore.org/documents/dcmi-terms/ (last visited 25th July 2012).
9. Dublin Core Metadata Initiative (2010). Dublin Core Metadata Element Set, Version 1.1. DCMI Recommendation. http://dublincore.org/documents/dces/ (last visited 25th July 2012).
10. Groza, T., Handschuh, S., Decker, S. (2011). Capturing Rhetoric and Argumentation Aspects within Scientific Publications. *Journal on Data Semantics*, 15: 1-36. doi:10.1007/978-3-642-22630-4_1.
11. Hammond, T. (2008). RDF Site Summary 1.0 Modules: PRISM. http://nurture.nature.com/rss/modules/mod_prism.html (last visited 25th July 2012).
12. Lagoze, C., Van de Sompel, H., Johnston, P., Nelson, M., Sanderson, R., Warner, S. (2008). ORE User Guide – Primer. Open Archives Initiative, Object Reuse and Exchange. http://www.openarchives.org/ore/1.0/primer (last visited 25th July 2012).
13. IFLA Study Group on the Functional Requirements for Bibliographic Records (2009). Functional Requirements for Bibliographic Records Final Report. International Federation of Library Associations and Institutions. http://www.ifla.org/files/cataloguing/frbr/frbr_2008.pdf (last visited 25th July 2012).
14. International Digital Enterprise Alliance (2009). Publishing Requirements for Industry Standard Metadata Specification Version 2.0. Alexandria, VA, USA: IDEAlliance. http://www.idealliance.org/specifications/prism (last visited 25th July 2012).
15. Jörg, B. (2008). Towards the Nature of Citations. In Poster Proceedings of the 5th International Conference on Formal Ontology in Information Systems (FOIS 2008). http://www.dfki.de/~brigitte/publications/FOIS08_Poster_BrigitteJoerg.pdf (last visited 25th July 2012).
16. Miles, A., Bechhofer, S. (2009). SKOS Simple Knowledge Organization System Reference. W3C Recommendation, 18 August 2009. World Wide Web Consortium. http://www.w3.org/TR/2009/REC-skos-reference-20090818/ (last visited 25th July 2012).
17. Peroni, S., Shotton, D., Vitali, F. (2012). Scholarly publishing and Linked Data: describing roles, statuses, temporal and contextual extents. To appear in Proceedings of the 8th International Conference on Semantic Systems (i-Semantics 2012). Postprint available at http://palindrom.es/phd/wp-content/uploads/2010/07/toc_isemantics_cr.pdf (last visited 25th July 2012).
18. Peroni, S., Shotton, D., Vitali, F. (2012). The Live OWL Documentation Environment: a tool for the automatic generation of ontology documentation. To appear in Proceedings of the 18th International Conference on Knowledge Engineering and Knowledge Management (EKAW 2012). Postprint available at http://palindrom.es/phd/wp-content/uploads/2010/07/lode_ekaw2012_cr.pdf (last visited 25th July 2012).
19. Rector, A. (2003). Modularisation of Domain Ontologies Implemented in Description Logics and related formalisms including OWL. In Proceedings of the 2nd international conference on Knowledge Capture (K-CAP 03): 121-128. doi:10.1145/945645.945664.
20. Schneider, J., Groza, T., Passant, A. (2012). A Review of Argumentation for the Social Semantic Web. To appear in *Semantic Web Journal*. Special Issue: Semantic Web – Interoperability, Usability, Applicability. Postprint available at http://www.semantic-web-journal.net/content/review-argumentation-social-semantic-web (last visited 25th July 2012).
21. Shotton, D. (2009). Semantic Publishing: the coming revolution in scientific journal publishing. *Learned Publishing* 22 (2): 85-94. doi:10.1087/2009202.
22. Shotton, D. (2010). CiTO, the Citation Typing Ontology. *Journal of Biomedical Semantics* 1 (1), S6. doi:10.1186/2041-1480-1-S1-S6.
23. Shotton, D., Caton, C., Klyne, G. (2010). Ontologies for sharing, ontologies for use. In The Ontogenesis Knowledge Blog. 2010:Paper 3. http://ontogenesis.knowledgeblog.org/2010/01/22/ontologies-for-sharing/ (last visited 25th July 2012).
24. Shotton, D., Portwin, K., Klyne, G., Miles, A. (2009). Adventures in Semantic Publishing: Exemplar Semantic Enhancements of a Research Article. *PLoS Computational Biology* 5 (4): e1000361. doi:10.1371/journal.pcbi.1000361.
25. Sure, Y., Bloehdorn, S., Haase, P., Hartmann, J., Oberle, D. (2005). The SWRC Ontology – Semantic Web for Research Communities. In Proceedings of the 12th Portuguese Conference on Artificial Intelligence (EPIA 2005): 218-231. doi:10.1007/11595014_22.
26. Teufel, S., Siddharthan, A., Tidhar, D. (2006). An annotation scheme for citation function. In Proceedings of the 7th SIGdial Workshop on Discourse and Dialogue (SigDIAL06): 80-87. http://dl.acm.org/citation.cfm?id=1654612 (last visited 25th July 2012).
27. Toulmin, S. (1959). The uses of argument. Cambridge, UK: Cambridge University Press. ISBN: 0521827485.

28. W3C OWL Working Group (2009). OWL 2 Web Ontology Language Document Overview. W3C Recommendation, 27 October 2009. World Wide Web Consortium. http://www.w3.org/TR/owl2-overview/ (last visited 25th July 2012).
29. de Waard, A. (2010). From Proteins to Fairytales: Directions in Semantic Publishing. *IEEE Intelligent Systems* 25 (2): 83-88. doi:10.1109/MIS.2010.49



Prefixes

rdfs:	http://www.w3.org/2000/01/rdf-schema#
fabio:	http://purl.org/spar/fabio/
frbr:	http://purl.org/vocab/frbr/core#


fabio:isSchemeOf only fabio:SubjectDiscipline
DisjointWith: fabio:DisciplineDictionary
fabio:isSchemeOf only fabio:SubjectTerm
rdfs:subClassOf
rdfs:subClassOf
fabio:DisciplineDictionary
rdfs:subClassOf
skos:ConceptScheme
rdfs:subClassOf
fabio:TermDictionary
skos:inScheme
fabio:hasDiscipline
skos:inScheme some fabio:TermDictionary
skos:Concept
skos:inScheme some fabio:DisciplineDictionary
rdfs:subClassOf
rdfs:subClassOf
rdfs:subClassOf
rdfs:subClassOf
fabio:SubjectTerm
fabio:SubjectDiscipline
fabio:hasSubjectTerm
fabio:hasDiscipline
frbr:Endeavour
Prefixes
rdfs: http://www.w3.org/2000/01/rdf-schema#
fabio: http://purl.org/spar/fabio/
skos: http://www.w3.org/2004/02/skos/core#

Clustering of CiTO relationships by similarity

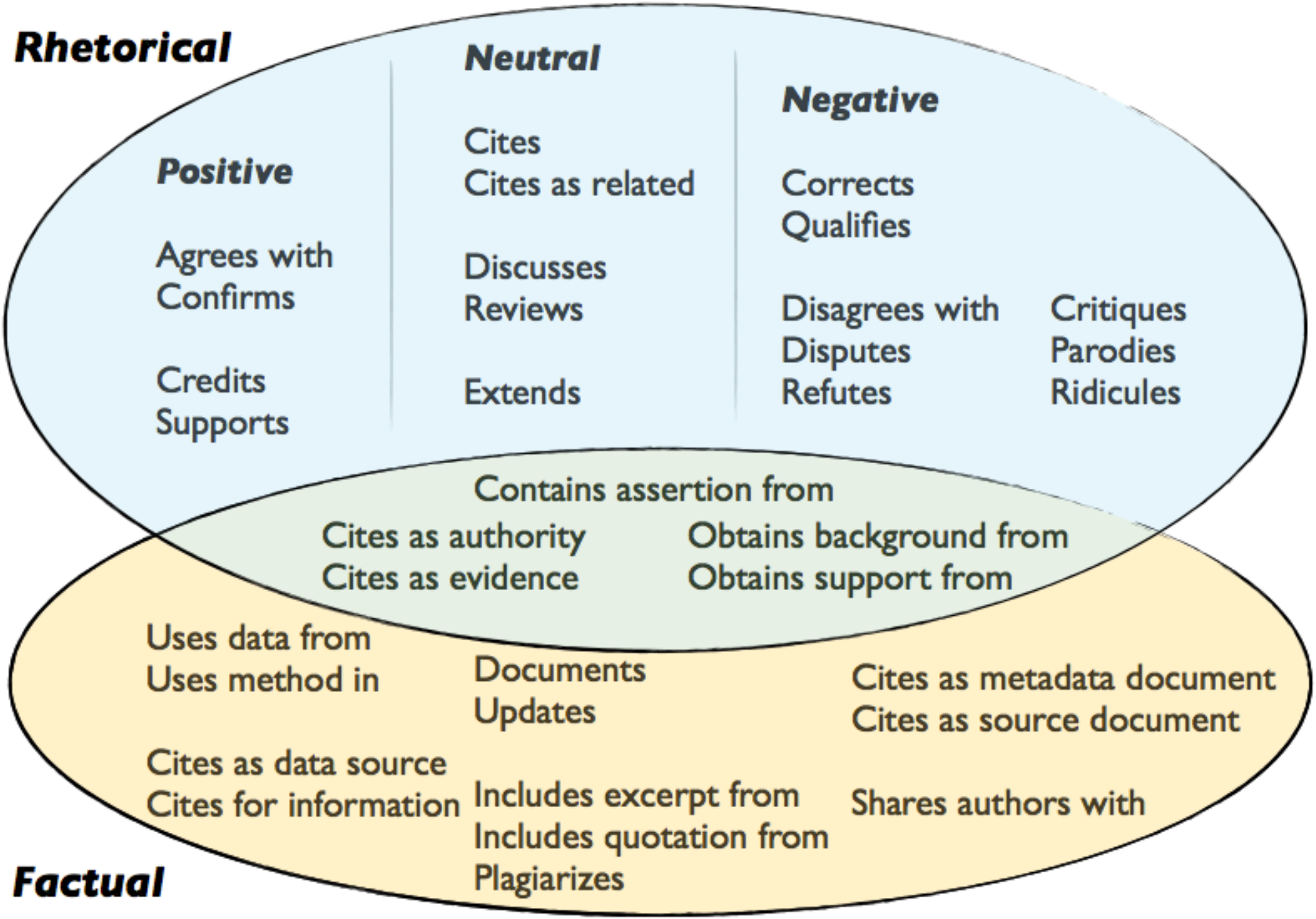


Figure 1. The main FRBR object properties relating FRBR endeavours (work, expression, manifestation, item), and the related new object properties introduced by FaBiO (fabio:hasManifestation, fabio:hasRepresentation, fabio:hasPortrayal) to provide shortcuts between Work and Manifestation, Work and Item, and Expression and Item, respectively.

Figure 2. The extension to the common SKOS classes and relations implemented in FaBiO.

Figure 3. The diagram shows the CiTO v 2.0 object properties grouped in terms of their characterisation as rhetorical and/or factual, and, for the former, in terms of their connotation (positive, neutral or negative). All the properties shown, except cito:cites and cito:sharesAuthorsWith, are sub-properties of cito:cites itself. The inverse property of *cito:cites*, namely *cito:isCitedBy,* and its inverse sub-properties, are not shown.