



Portal Ontology

Semantic Web Community Portal Project

Holger Lausen, Michael Stollberg

DERI Research Report

January 2004

Deliverable:
D1

Version:
0.2

Date:
19-Jan-04

DERI Ireland

University Road

Galway

IRELAND

www.deri.ie

DERI Austria

Technikerstrasse 13

Table of Contents

1 EXECUTIVE SUMMARY	1
1.1 Project Members	1
2 DESIGN OBJECTIVES.....	2
2.1 Purpose of Use	2
2.2 General Structure	2
3 ONTOLOGY DETAIL DESCRIPTION.....	3
4 OPEN POINTS.....	5
4.1 Status of Development	5
4.2 Proposal for Different Approach.....	5



2 Design Objectives

2.1 Purpose of Use

This ontology shall capture common sense knowledge and enable interoperability. It should not be treated as static but as starting point for a community driven evolution.

This ontology focuses in the current version on a domain ontology describing scientific publications. We choose the concept “Publication” as starting point, since scientific publication will be one of the aspects of the semanticweb.org portal and there exists already a well known and established format for its annotation: BibTex.

Other aspects like application specific ones are currently not modeled, but may be objective of future efforts. For example another ontology could describe application specific aspects like different means of presenting information (pages, cells, dropdown lists, etc.). The combination of both ontology could provide a declarative specification of the portal.

2.2 General Structure

The presented ontology consists of one Root Concept “Publication” and 16 sub concepts according to the BibTex Entry Types defined in [Patashink 88]:

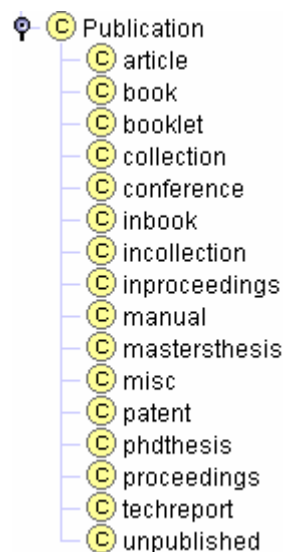


Figure 1: BibTex Entry Types

These Concepts were introduced in 1985 and are meanwhile widely used and accepted². BibTeX is a program and file format designed by Oren Patashnik and Leslie Lamport in 1985 for the [LaTeX](#)³ document preparation system. Detailed rationalization for the concepts can be found in the literature [Patashnik 88] [Lamport 86]. In the following we will have a more detailed look at how this file format can be translated into an ontology with concepts and properties.

3 Ontology Description

3.1 General Comments

3.1.1 Existing BibTeX Ontology

There are currently two public ontologies based on the bibtex format:

- **DARPA/DAML BibTeX Ontology** (2001)
<http://www.daml.org/ontologies/115>
- **KAON BibTeX Ontology**
http://kaon.semanticweb.org/ontos/bibtex.kaon/ontology_view

The DARPA/DAML BibTeX ontology is a one-to-one mapping to the BibTeX types and fields, as defined in [Goosens]. But it is only an enumeration of classes and properties with comments. There exist no local restrictions for the properties on the classes, even though [Patashnik 88] defines them. Also the range of the properties is not defined. This ontology was taken as starting point and converted to OWL and restrictions were added.

The KAON BibTeX Ontology is loosely based on BibTeX (some concepts removed) but EU Project specific (e.g. Concept Deliverable) were added. This Ontology is only available in the KAON format and could therefore not be examined in detail. Also no related documentation was published.

3.1.2 Cardinality Restrictions

[Patashnik 88] implies some cardinality and domain restrictions. All attributes defined in [Lamport 86] are classified into: required, optional and ignored. This is in our ontology translated to “hasCardinality=1”, maxCardinality=1 and no statement.

² Example of pages providing bibliographic information as Bibtex Entry: <http://citeseer.nj.nec.com/>, <http://portal.acm.org/>

³ <http://tex.loria.fr/english/index.html>



3.1.3 Property ranges

For all property ranges we define “xsd:String” as range, since this is the definition naturally implies by [Patashink 88]. In A later stage we might want to alter them to allow to reference to existing instances (e.g. author).

3.2 Example Concepts

We now present the Concept Article. This Concept is for annotation for Articles from Journals and Magazines only. It requires the following fields: author, journal, title, year. Optionally volume, number, pages, month and note may be given. For the readers convenience the Annex A contains a copy of the relevant section in [Patashink 88].

This is how it translates into out ontology:

3.3 Class article

Concrete Class Extends [Publication](#)

Direct Subclasses:

Class Documentation: An article from a journal or magazine.

Slot name	Documentation	Type	Cardinality
<i>title</i>	The work's title, typed as explained in the LaTeX book. [@@to be looked up]	String	0:1
<i>note</i>	Any additional information that can help the reader. The first word should be capitalized.	String	0:1
<i>journal</i>	A journal name. Abbreviations are provided for many journals.	String	0:1
<i>number</i>	The number of a journal, magazine, technical report, or of a work in a series. An issue of a journal or magazine is usually identified by its volume and number; the organization that issues a technical report usually gives it a number; and sometimes books are given numbers in a named series.	String	0:1

<i>year</i>		String	0:1
<i>pages</i>	One or more page numbers or range of numbers, such as 42--111 or 7,41,73--97 or 43+ (the '+' in this last example indicates pages following that don't form a simple range). To make it easier to maintain Scribe-compatible databases, the standard styles convert a single dash (as in 7-33) to the double dash used in TeX to denote number ranges (as in 7--33).	String	0:1
<i>volume</i>	The volume of a journal or multi-volume book.	String	0:1
<i>author</i>	The name(s) of the author(s), in the format described in the LaTeX book. [@@to be looked up]	String	0:1
<i>month</i>	The month in which the work was published or, for an unpublished work, in which it was written. You should use the standard three-letter abbreviation, as described in Appendix B.1.3 of the LaTeX book.	String	0:1

[@@to be extended] If this is accepted by the working group as a feasible documentation and choice of domain the document will be extended.

4 Open Points

4.1 Status of Development

- the ontology needs discussion / finding a consensus throughout the project members
- the current status must be seen as initial proposal

5 Development Notes

Representation Language for the Ontology is currently OWL. The concrete subset to be used is not yet defined.



5.1 Development Tool

The Ontology is developed with Protégé (Beta-Build 111, OWL & ezOWL Plugin). It is strongly recommended to download the latest version from protégé.stanford.edu, since current updates are significantly improving usability.

5.2 Multi User Support and Versioning

Currently we do not yet have both in place. Both should be agreed on soon, to facilitate the consensus building process. The Ontology is not to be intended stable in this phase of development.

6 TO DO

- Versioning / multi user development concept
- Feedback

7 References

[Patashink 88] Oren Patashnik : BIBTEXing, February 8, 1988,
<http://tex.loria.fr/bibdex/btxdoc.pdf>

[Lamport 86] Leslie Lamport: *LaTeX: A Document Preparation System*, 1986, Addison-Wesley

[Goossens]M. Goossens, F.Mittelbach and Samarin: The LaTeX Companion

Annex A

Types

article An article from a journal or magazine. Required _elds: author, title, journal, year. Optional _elds: volume, number, pages, month, note.

book A book with an explicit publisher. Required _elds: author or editor, title, publisher, year. Optional _elds: volume or number, series, address, edition, month, note.

booklet A work that is printed and bound, but without a named publisher or sponsoring institution. Required _eld: title. Optional _elds: author, howpublished, address, month, year, note.

conference The same as INPROCEEDINGS, included for *Scribe* compatibility.

inbook A part of a book, which may be a chapter (or section or whatever) and/or a range of pages. Required _elds: author or editor, title, chapter and/or pages, publisher, year. Optional _elds: volume or number, series, type, address, edition, month, note.



incollection A part of a book having its own title. Required _elds: author, title, booktitle, publisher, year. Optional _elds: editor, volume or number, series, type, chapter, pages, address, edition, month, note.

inproceedings An article in a conference proceedings. Required _elds: author, title, booktitle, year. Optional _elds: editor, volume or number, series, pages, address, month, organization, publisher, note.

manual Technical documentation. Required _eld: title. Optional _elds: author, organization, address, edition, month, year, note.

mastersthesis A Master's thesis. Required _elds: author, title, school, year. Optional _elds: type, address, month, note.

misc Use this type when nothing else _ts. Required _elds: none. Optional _elds: author, title, howpublished, month, year, note.

phdthesis A PhD thesis. Required _elds: author, title, school, year. Optional _elds: type, address, month, note.

proceedings The proceedings of a conference. Required _elds: title, year. Optional _elds: editor, volume or number, series, address, month, organization, publisher, note.

techreport A report published by a school or other institution, usually numbered within a series. Required _elds: author, title, institution, year. Optional _elds: type, number, address, month, note.

unpublished A document having an author and title, but not formally published. Required _elds: author, title, note. Optional _elds: month, year.

Fields

address Usually the address of the publisher or other type of institution. For major publishing houses, van Leunen recommends omitting the information entirely. For small publishers, on the other hand, you can help the reader by giving the complete address.

annote An annotation. It is not used by the standard bibliography styles, but may be used by others that produce an annotated bibliography.

author The name(s) of the author(s), in the format described in the LATEX book.

booktitle Title of a book, part of which is being cited. See the LATEX book for how to type titles. For book entries, use the title _eld instead.

chapter A chapter (or section or whatever) number.

crossref The database key of the entry being cross referenced.

edition The edition of a book|for example, \Second". This should be an ordinal, and should have the _rst letter capitalized, as shown here; the standard styles convert to lower case when necessary.

editor Name(s) of editor(s), typed as indicated in the LATEX book. If there is also an author _eld, then the editor _eld gives the editor of the book or collection in which the reference appears.



howpublished How something strange has been published. The `_rst` word should be capitalized.

institution The sponsoring institution of a technical report.

journal A journal name. Abbreviations are provided for many journals; see the *Local Guide*.

key Used for alphabetizing, cross referencing, and creating a label when the `\author` information (described in Section 4) is missing. This `_eld` should not be confused with the key that appears in the `\cite` command and at the beginning of the database entry.

month The month in which the work was published or, for an unpublished work, in which it was written. You should use the standard three-letter abbreviation, as described in Appendix B.1.3 of the LATEX book.

note Any additional information that can help the reader. The `_rst` word should be capitalized.

number The number of a journal, magazine, technical report, or of a work in a series. An issue of a journal or magazine is usually identified by its volume and number; the organization that issues a technical report usually gives it a number; and sometimes books are given numbers in a named series.

organization The organization that sponsors a conference or that publishes a manual.

pages One or more page numbers or range of numbers, such as 42--111 or 7,41,73--97 or 43+ (the '+' in this last example indicates pages following that don't form a simple range). To make it easier to maintain *Scribe*-compatible databases, the standard styles convert a single dash (as in 7-33) to the double dash used in TEX to denote number ranges (as in 7--33).

publisher The publisher's name.

school The name of the school where a thesis was written.

series The name of a series or set of books. When citing an entire book, the `title_eld` gives its title and an optional `series_eld` gives the name of a series or multi-volume set in which the book is published.

title The work's title, typed as explained in the LATEX book.

type The type of a technical report|for example, `\Research Note`".

volume The volume of a journal or multivolume book.

year The year of publication or, for an unpublished work, the year it was written. Generally it should consist of four numerals, such as 1984, although the standard styles can handle any year whose last four nonpunctuation characters are numerals, such as `\(about 1984)`'.

