

Knowledge Organization Systems, Simple

Extending SKOS to support Simple Knowledge Organization Systems

Simon Spero, (ses@unc.edu)

School of Information and Library Science
University of North Carolina At Chapel Hill

Abstract

This document describes the Knowledge Organization System, Simple (KOSS), a set of extensions to SKOS that models the semantics of traditional KOS. Where possible, KOSS classes and properties are defined as extensions of corresponding SKOS features, providing backwards-compatibility for legacy SKOS applications.

1 Introduction

”The word butterfly is not a real butterfly. There is the word and there is the butterfly. If you confuse these two items people have the right to laugh at you.”

Leonard Cohen - How to speak poetry.

Although they share many similar features, Classes in modelling languages like OWL, Subjects in Knowledge Organization Systems such as the LCSH, and SKOS Concepts are different types of things. Classes describe sets of things. For example, the class ”Butterflies” refers to things that are a butterfly. Subjects, in contrast, describe what something is about; for example, the subject ”Butterflies” refers to documents that are about butterflies.

In the initial public draft of SKOS,¹ Concepts were modeled after traditional Subjects; in 2008 the model was changed to loosen this restriction. This broadening of the scope may have removed certain desirable properties that are useful for modelling subject languages in the semantic web. This document describes some simple specializations of SKOS that may be useful for modelling KOS systems such as the LCSH.

2 Knowledge Organization Systems

See e.g. Z39.19² and Svenonius.³ I will be using using standard relationship notations such as RT, BT/NT, BTG/NTG, BTI/NTI, and BTP/NTP . The relationships defined are intended to have the standard meanings

¹Alistair Miles and Dan Brickley (2005). *SKOS Core Vocabulary Specification W3C Working Draft - November 2nd 2005*. Draft Specification. W3C. URL: <http://www.w3.org/TR/2005/WD-swbp-skos-core-spec-20051102>.

²NISO Z39.19 (2005). *Guidelines for the Construction, Format, and Management of Monolingual Controlled Vocabularies*. National Information Standards Organization. URL: <http://www.niso.org/standards/resources/z39-19-2005.pdf>.

³Elaine Svenonius (2000). *The Intellectual Foundation of Information Organization*. Cambridge, Mass.: MIT Press. ISBN: 0262194333 (hc : alk. paper). URL: <http://www.netlibrary.com/AccessProduct.aspx?ProductId=39954>.

2.1 Types of Knowledge Organization Systems

2.2 Terns, Subjects and Concepts

2.3 Relationships in Knowledge Organization Systems

3 SKOS

There are several issues in using SKOS⁴ to model traditional Knowledge Organization Systems.

1. Concepts aren't defined in terms of subjects.
2. Does not support subdivided subjects.
3. Does not support pre-combined subjects.
4. Does not support assertion of hierarchical relationships
 - (a) (broader is not transitive
 - (b) broaderTransitive is not supposed to be asserted
 - (c) "the semantics of skos:broader, [...] cannot enforce such transitivity"⁵

4 Brief introduction to the Semantic Web

TimBL⁶

4.1 RDF, Triples, and Linked Data

RDF⁷

4.2 OWL, the Web Ontology Language

OWL⁸

4.3 SWRL, the Semantic Web Rules Language

SWRL⁹

⁴Alistair Miles and Sean Bechhofer (2009). *SKOS Simple Knowledge Organization System Reference*. Proposed Recommendation. W3C. URL: <http://kent.w3.org/TR/skos-reference>.

⁵Antoine Isaac and Ed Summers (2009). *SKOS Primer*. Working Draft. W3C. URL: <http://www.w3.org/TR/skos-primer>.

⁶Tim Berners-Lee, James Hendler, and Ora Lassila (2001). "The Semantic Web". In: *Scientific American* (May 2001). URL: <http://www.sciam.com/article.cfm?articleID=00048144-10D2-1C70-84A9809EC588EF21&catID=2>.

⁷David Beckett (2004). *RDF/XML Syntax Specification (Revised)*. Tech. rep. World Wide Web Consortium. URL: <http://www.w3.org/TR/2004/REC-rdf-syntax-grammar-20040210/>; W3C. *Resource Description Framework (RDF)/W3C Semantic Web Activity*. URL: <http://www.w3.org/RDF/>.

⁸OWL Working Group (2009). *OWL 2 Web Ontology Language Document Overview*. Working Draft. W3C. URL: <http://www.w3.org/TR/owl2-overview/>.

⁹I. Horrocks et al. (2004). *SWRL: A semantic web rule language combining OWL and RuleML*. Member Submission. World Wide Web Consortium. URL: <http://www.w3.org/Submission/SWRL/>.

5 Using Controlled Natural Language to define Ontologies

5.1 Usability Problems of OWL and SWRL for Untrained Users

Many people in the Knowledge Organization and Library and Information Science communities do not have a background in computer science or mathematical logic. This can make it hard for members of those communities to understand the details of Semantic Web specifications, and to use those specifications to create their own ontologies.¹⁰

One approach to dealing with this issue is to use a Controlled Natural Language to specify ontologies. Controlled Natural Languages are subsets of human languages which are simplified to reduce ambiguity and to allow for easier processing by computers. Several languages have developed specifically to support the development of ontologies, rule sets, and related forms of knowledge representation.¹¹

5.2 Attempto Controlled English

Attempto Controlled English (ACE) is a controlled version of English that has been developed by the Attempto project at the University of Zurich. ACE was adopted by the REWERSE research Network of Excellence on "Reasoning on the Web",¹² and has well defined mappings from and to OWL and SWRL.¹³ In this document, all specifications will be made using Attempto Controlled English (ACE).¹⁴

5.3 Examples

These sentences in ACE are equivalent to the OWL ontology in Figure 1.

```
Every mammal is an animal.  
Every weasel is a mammal.  
Everything that a weasel eats is an egg.
```

¹⁰Alan. Rector et al. (2004). "Owl pizzas: Practical experience of teaching owl-dl: Common errors and common patterns". In: *Engineering Knowledge in the Age of the Semantic Web*. Lecture Notes in Computer Science. Springer, pp. 63–81.

¹¹Paul R Smart (2008). *Controlled Natural Languages and the Semantic Web*. Tech. rep. ITA/P12/SemWebCNL. School of Electronics and Computer Science, University of Southampton. URL: <http://eprints.ecs.soton.ac.uk/15735/>.

¹²Norbert E. Fuchs et al. (2006). *Attempto Controlled English and the Semantic Web*. Deliverable I2-D7. Network of Excellence on Reasoning on the Web with Rules and Semantics (REWERSE). URL: <http://rewerse.net/deliverables/m24/i2-d7.pdf>.

¹³Kaarel Kaljurand (2007). "Attempto Controlled English as a Semantic Web Language". PhD thesis. Faculty of MathematicsComputer Science, University of Tartu; Kaarel Kaljurand (2008). "ACE View — An Ontology and Rule Editor based on Controlled English". In: *Proceedings of the Poster and Demonstration Session at the 7th International Semantic Web Conference (ISWC2008)*. Ed. by Christian Bizer and Anupam Joshi. Vol. 401. CEUR Workshop Proceedings.

¹⁴To simplify some descriptions, I have gone slightly beyond ACE by allowing variable to be used as classes. Since the classes in question are not in the KOSS ontology, but instead in the ontology we are inferring from the knowledge organization system that we are using KOSS to describe, this punning is mostly harmless

```

Ontology(
  http://attempto.ifi.uzh.ch/ontologies/owlswrl/test
  SubClassOf(
    Class(:mammal)
    Class(:animal)
  )
  SubClassOf(
    Class(:weasel)
    Class(:mammal)
  )
  SubClassOf(
    ObjectIntersectionOf(
      Class(owl:Thing)
      ObjectSomeValuesFrom(
        InverseObjectProperty(ObjectProperty(:eat))
        Class(:weasel)
      )
    )
    Class(:egg)
  )
)

```

Figure 1: The Weasel Ontology

6 KOSS Spec

6.1 Subjects and Documents

Every subject is a skos:Concept that something is-about.
Everything that something is-about is a subject.
Everything that is-about something is a document.

These definitions explicitly define subjects are what things are about. Following M.K. Buckland (1997). “What is a “Document”?” In: *Journal of the American Society for Information Science* 48.9, pp. 804–809. URL: <http://people.ischool.berkeley.edu/~buckland/whatdoc.html> and Suzanne Briet (1951). *What is Documentation?* Ed. by Ronald E Day and Laurent Martinet. Paris, France: Édit. URL: <http://ella.slis.indiana.edu/~roday/briet.htm> we treat anything that is about something as being a document.

6.2 Associative relationship: RT

If X RT Y then X skos:related Y.
If X RT Y then Y RT X.
If something RT something then it is a subject.

The RT relationship is basically the same as skos 'related' ; it is introduced for completeness.

6.3 The Hierarchical Relationship: BT and NT

If X BT Y then X skos:broader Y.
If X NT Y then X skos:narrower Y.
If X BT Y then Y NT X.

If something is-about a subject that BT Y then it is-about Y.

If X BT something that BT Y then X BT Y.

If something BT something then it is a subject.

The definition of the BT relationship denotes the traditional Hierarchical relationship. It differs from the skos 'broader' relationship in a number of key respects. Firstly, it specifies that documents that are about a narrower term are also about the broader term; secondly, it defines the relationship as being necessarily transitive.

6.4 Generic relationships: BTG and NTG

If X BTG Y then X BT Y.
If X NTG Y then X NT Y.

IF X BTG Y then Y NTG X.

If X BTG something that BTG Y then X BTG Y.

Everything that BTG something is a generic-subject.

If a subject that has-owl-class C

```

    BTG a subject that has-owl-class D
    then every C is a D .

```

Here we define BTG as a sub-relationship of BT; it can be used to express a relationship between two subjects that arises due an underlying genus/species relationship between two classes; for example, the subject "Weasels" BTG "Mammals" because weasels are mammals.

Note that BTG is a transitive. If there is a chain of subjects, each of which is BTG to the next, then each narrower subject is BTG to all the broader ones. If any of the links are not BTG but some other kind of hierarchical relationship, the chain is broken, and only BT is inferred.

6.5 Instantive relationships: BTI and NTI

```

If X BTI Y then X BT Y.

```

```

If X NTI Y then X NT Y.

```

```

If X BTI Y then Y NTI X.

```

```

If X BTI something that BTG Y then X BTI Y.

```

```

If X NTG something that NTI Y then X NTI Y.

```

```

Everything that BTI something is a named-individual-subject.

```

```

If a subject that has-individual I
    BTI a subject that has-owl-class C
    then I is a C.

```

BTI can be used to define relationships between subjects related to a class, and subjects related to specific members of that class. For example "Andes" BTI "Mountain Ranges" because the Andes are a mountain range.

The BTI relationship is not transitive, but does follow BTG links. For example, if "Wally the weasel" BTI "Weasels" and "Weasels" BTG "Mammals", "Wally the weasel" BTI "Mammals".

6.6 Partitive Relationships: BTP/NTP

```

If X BTP Y then X BT Y.

```

```

If X NTP Y then X NT Y.

```

```

If X BTP Y then Y NTP X.

```

```

If X BTP something that BTP Y then X BTP Y.

```

```

If X has-part Y then P part-of X.

```

```

If X part-of Y then X has-part Y.

```

```

/* Not real ACE; need to create a restriction

```

```

    If a subject Part BTP a subject Whole and
        Part has-owl-class Part-Class and
        Whole has-owl-class Whole-Class
    then
        Whole-Class has-part at least 1 thing that is a Part-Class.
*/

```

BTP is used to express relationships between parts and wholes. It only applies to subjects describing parts that are necessarily part of a whole.

For example, "Car Steering Wheels" BTP "Cars".

This relationship introduces a cardinality restriction on constraint on has-part for the underlying Classes associated with the subjects.

6.7 Subdivisions

Every subdivision is something that is
a class-changing-subdivision or that is
a class-preserving-subdivision .

Every chronological-subdivision is a class-preserving-subdivision.
Every geographic-subdivision is a class-preserving-subdivision.
Every form-genre-subdivision is a class-changing-subdivision .
Every topical-subdivision is a class-changing-subdivision.

If a subject S has-main-subject M then S BT M.

If a subject S is subdivided by a subdivision
that has-associated-subject A
then S BT A.

If a subject S has-main-subject M
and S is subdivided by a class-preserving-subdivision
then S BTG M.

If a subject S is subdivided by a class-changing-subdivision C
and C has-associated-subject A
then S BTG A.

If a document is-about a subject that is
subdivided by a form-genre-subdivision
and the form-genre-subdivision has-associated-class C then
the document is a C.

We divide subdivisions into two basic types; those that preserve the class of the main subject they divide, and those which change it. Both kinds of subdivision introduce BT relationships from the subdivided subject to the main subject, and to the subject associated with the subdivision. Form/genre headings introduce an additional complication; they describe properties of documents that are about the subject. Thus something that is about "Weasels-Motion pictures" is a motion picture.

Class preserving subdivisions introduce a BTG relationship to the main subject; class changing subdivisions to the subject associated with the subdivision.

References

Beckett, David (2004). *RDF/XML Syntax Specification (Revised)*. Tech. rep. World Wide Web Consortium.
URL: <http://www.w3.org/TR/2004/REC-rdf-syntax-grammar-20040210/>.

- Berners-Lee, Tim, James Hendler, and Ora Lassila (2001). “The Semantic Web”. In: *Scientific American* (May 2001). URL: <http://www.sciam.com/article.cfm?articleID=00048144-10D2-1C70-84A9809EC588EF21&catID=2>.
- Briet, Suzanne (1951). *What is Documentation?* Ed. by Ronald E Day and Laurent Martinet. Paris, France: Édit. URL: <http://ella.slis.indiana.edu/~roday/briet.htm>.
- Buckland, M.K. (1997). “What is a “Document”?” In: *Journal of the American Society for Information Science* 48.9, pp. 804–809. URL: <http://people.ischool.berkeley.edu/~buckland/whatdoc.html>.
- Fuchs, Norbert E. et al. (2006). *Attempto Controlled English and the Semantic Web*. Deliverable I2-D7. Network of Excellence on Reasoning on the Web with Rules and Semantics (REWERSE). URL: <http://rewerse.net/deliverables/m24/i2-d7.pdf>.
- Horrocks, I. et al. (2004). *SWRL: A semantic web rule language combining OWL and RuleML*. Member Submission. World Wide Web Consortium. URL: <http://www.w3.org/Submission/SWRL/>.
- Isaac, Antoine and Ed Summers (2009). *SKOS Primer*. Working Draft. W3C. URL: <http://www.w3.org/TR/skos-primer>.
- Kaljurand, Kaarel (2007). “Attempto Controlled English as a Semantic Web Language”. PhD thesis. Faculty of Mathematics Computer Science, University of Tartu.
- (2008). “ACE View — An Ontology and Rule Editor based on Controlled English”. In: *Proceedings of the Poster and Demonstration Session at the 7th International Semantic Web Conference (ISWC2008)*. Ed. by Christian Bizer and Anupam Joshi. Vol. 401. CEUR Workshop Proceedings.
- Miles, Alistair and Sean Bechhofer (2009). *SKOS Simple Knowledge Organization System Reference*. Proposed Recommendation. W3C. URL: <http://kent.w3.org/TR/skos-reference>.
- Miles, Alistair and Dan Brickley (2005). *SKOS Core Vocabulary Specification W3C Working Draft - November 2nd 2005*. Draft Specification. W3C. URL: <http://www.w3.org/TR/2005/WD-swbpskos-core-spec-20051102>.
- NISO Z39.19 (2005). *Guidelines for the Construction, Format, and Management of Monolingual Controlled Vocabularies*. National Information Standards Organization. URL: <http://www.niso.org/standards/resources/Z39-19-2005.pdf>.
- OWL Working Group (2009). *OWL 2 Web Ontology Language Document Overview*. Working Draft. W3C. URL: <http://www.w3.org/TR/owl2-overview/>.
- Rector, Alan. et al. (2004). “Owl pizzas: Practical experience of teaching owl-dl: Common errors and common patterns”. In: *Engineering Knowledge in the Age of the Semantic Web*. Lecture Notes in Computer Science. Springer, pp. 63–81.
- Smart, Paul R (2008). *Controlled Natural Languages and the Semantic Web*. Tech. rep. ITA/P12/SemWebCNL. School of Electronics and Computer Science, University of Southampton. URL: <http://eprints.ecs.soton.ac.uk/15735/>.
- Svenonius, Elaine (2000). *The Intellectual Foundation of Information Organization*. Cambridge, Mass.: MIT Press. ISBN: 0262194333 (hc : alk. paper). URL: <http://www.netlibrary.com/AccessProduct.aspx?ProductId=39954>.
- W3C. *Resource Description Framework (RDF) / W3C Semantic Web Activity*. URL: <http://www.w3.org/RDF/>.