

RELcat: a Relation Registry for ISOcat data categories

Menzo Windhouwer

Max-Planck-Institute for Psycholinguistics

The Language Archive

Menzo.Windhouwer@mpi.nl

The ISOcat Data Category Registry

MPI/TLA and ISO TC 37 have developed ISOcat to provide a reusable set of (standardized) data categories (ISO 12620:2009).

A data category is an elementary descriptor in a linguistic structure or an annotation scheme.

ISOcat contains a shallow list of data categories, i.e., the registry doesn't store ontological relationships between data categories, as these relations are often application or even user specific.

www.isocat.org

A companion registry: RELcat

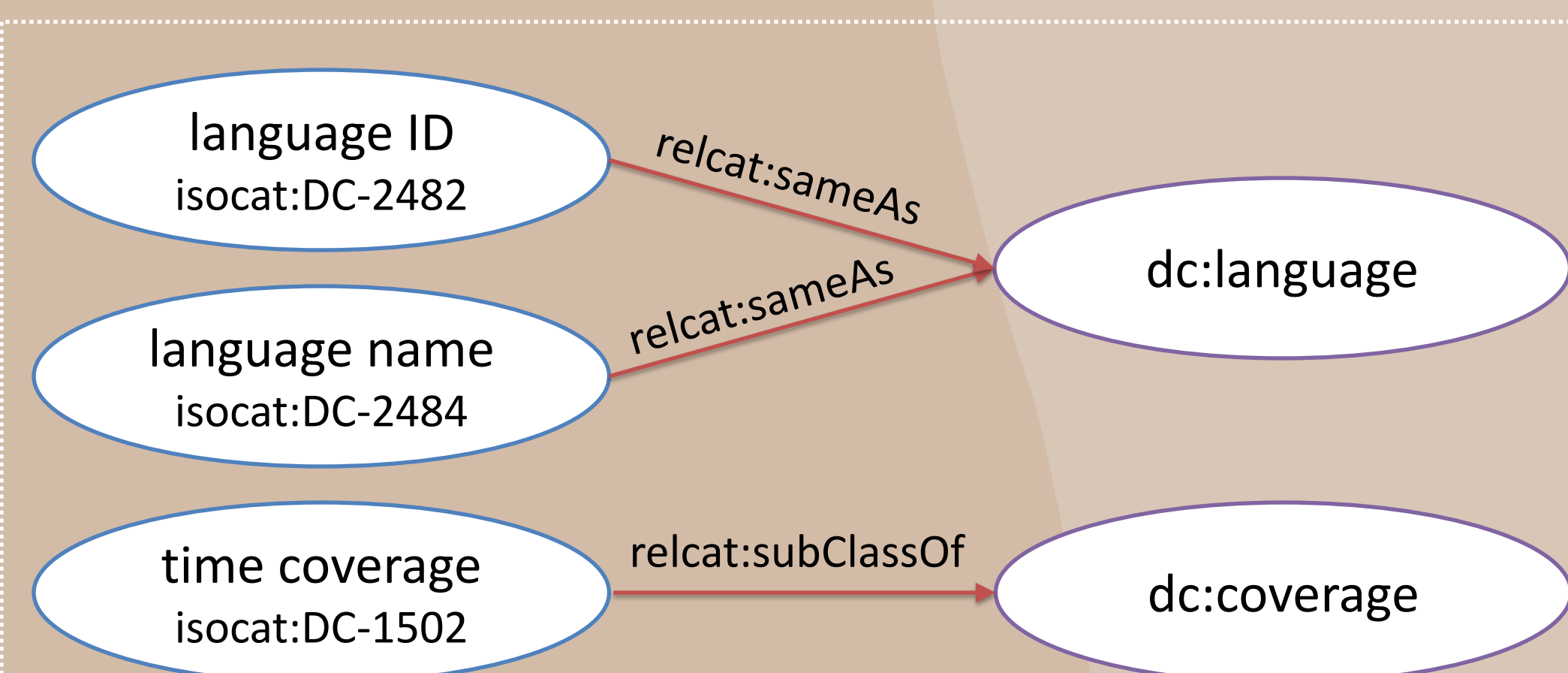
RELcat provides (groups of) users the ability to specify their own ontological relationships among ISOcat data categories but also to categories or concepts in other registries.

These sets of relations can be combined to form semantic networks, which can, for example, be exploited to find interesting resources based on their semantic proximity. These resources then need to be associated with ISOcat data categories.

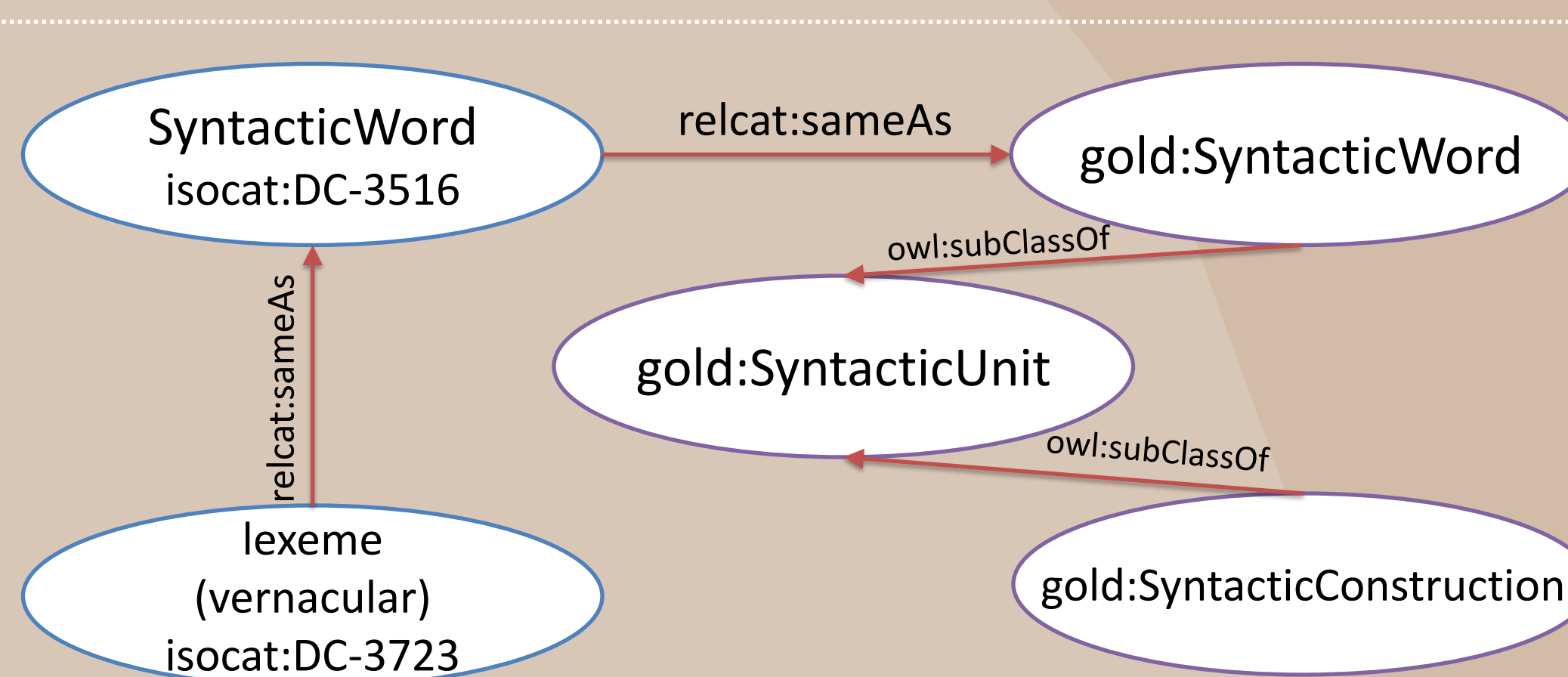
lux13.mpi.nl/relcat/

(alpha version)

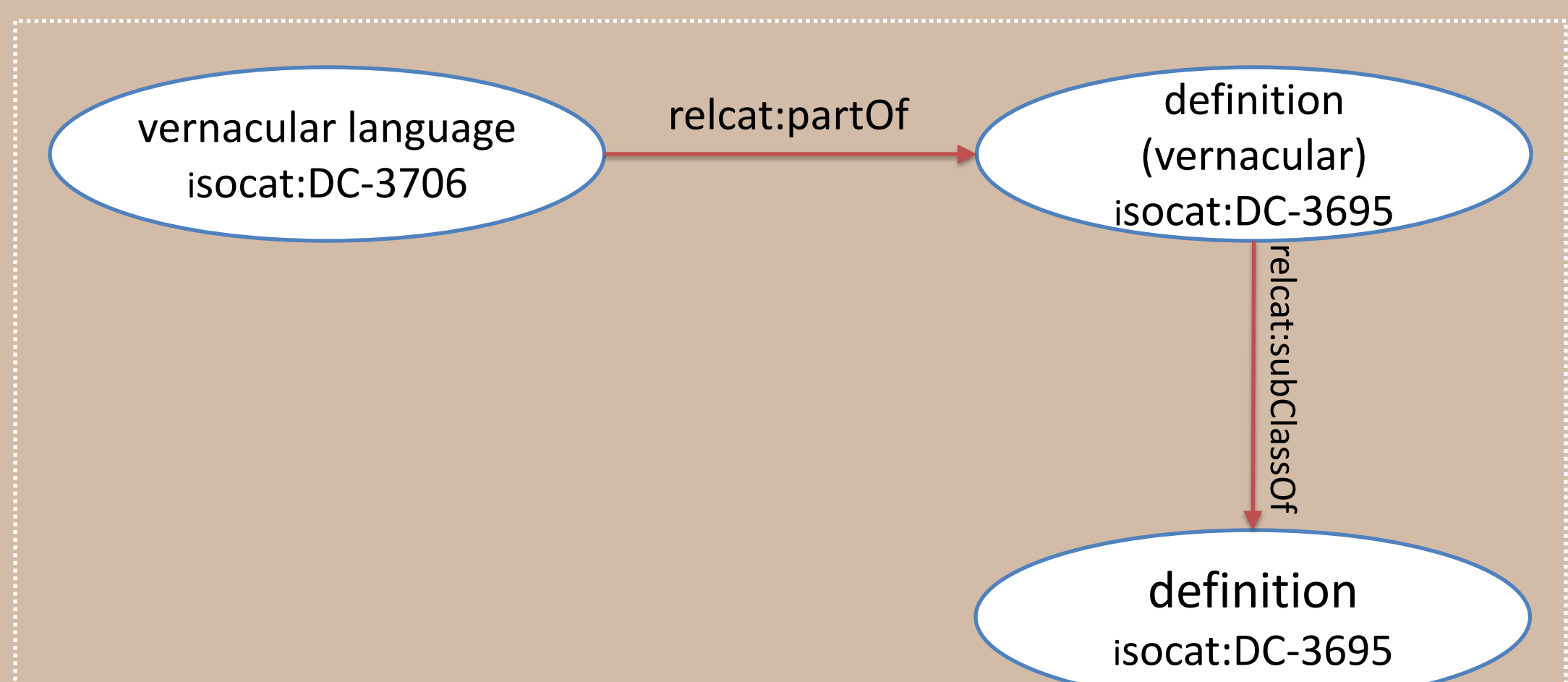
Examples taken from various relation sets



The equivalence of Component Metadata elements to Dublin Core elements used by the Semantic Mapping Component of CMD Infrastructure



A crosswalk to the GOLD ontology. Hooking up data categories to this set will enable generalization and specialization based on the GOLD subsumption hierarchy.



Granularity issues can be addressed by relating complex categories to their constituent categories or concepts using part-whole relationships.

Taxonomy of relation types

- **relcat:related**
 - **relcat:sameAs**
 - **owl:sameAs**
 - **owl:equivalentClass**
 - **owl:equivalentProperty**

A core set of relation types is available in an extensible taxonomy. Types from other vocabularies, e.g., OWL and SKOS, can be attached to this taxonomy. This allows generic algorithms to access all relations via the upper taxonomy, while vocabulary specific algorithms, e.g., OWL reasoners, can use specific extensions on unions of sets.

Implementation

An alpha version is available and provides read-only access to a growing number of relation sets. A web user interface to enable users to manage their own sets is under development. A set of web services to access and query the sets is available, and in use by the CLARIN infrastructure.

The implementation leverages Semantic Web technology, e.g., an RDF quad store, RDF-Plus reasoning and SPARQL queries.