Enrich bibliographies with metadata about individual bibliographic reference and identify authors motivations for citations. Each reference is treated as an individual, first-class entity, which can be accessed, filtered and grouped with other references according to different criteria ('coloring schemes').

**Background**

- **References**
  - Cited Entities

- **Citation Context**

- **Outgoing Citations**

- **Incoming Citations**

- **Citing Entities**

- **Which Citation Functions?**

**Methods & Technologies**

- **Machine Learning**
  - Supervised Approach

- **Semantic Web**
  - RDF, SPARQL, SPAR

- **NLP**
  - Tokenization, POS Tagging and Parsing, Corpora Analysis, Words Frequency and Similarity

**Project Timeline**

1. Analysis of Existing Citation Functions Schemas
2. Analysis of Criteria for Characterizing Citations
3. Build RDF Dataset on Test Articles
5. Development of the Citation Extraction Tool: Citation Network and Textual Citation Context
6. Machine Learning Classifier
7. Build an Annotated Corpus + Guidelines for Annotation
8. Users Evaluation
9. Validation

**Pipeline**

- **SCAR Exporter + API**

**Validation**

Result of 3-fold cross validation using ground-truth dataset of the citation functions classifier.

**Annotated Corpus**

- Annotated corpus about citation functions used to improve the performance of the classifier - 820 annotations (in progress)
- Ad-hoc guidelines in order to help annotators
- The annotated corpus will be soon released and freely available

**SCAR Citation Functions Schema**

- **Uses Method In**
  - The work of the citing entity uses a method presented in the cited entity
  - Amber force field (34) was used to model the DNA

- **Extends**
  - The citing entity extends, improves or continues the work of the cited entity
  - Our web application implements a Parallel Exploration paradigm which extends Parallel Focused Browsing (8)

- **Uses Data From**
  - The citing entity uses data presented in the cited entity
  - We mainly experiment on the Semantic 3D dataset (49)

- **Cites As Review**
  - The citing entity cites a work that reviews a series of works
  - Following previous studies (see Lin’s review (4)), we measured social capital […]

**Goal**

Enrich bibliographies with metadata about individual bibliographic reference and identify authors motivations for citations. Each reference is treated as an individual, first-class entity, which can be accessed, filtered and grouped with other references according to different criteria ('coloring schemes').

**Analysis of Existing Citation Functions Schemas**

**Development of a Proper Schema, Verified via Users Survey**

**Analysis of Criteria for Characterizing Citations**

**Build RDF Dataset on Test Articles**

**Development of the Citation Extraction Tool: Citation Network and Textual Citation Context**

**Machine Learning Classifier**

**Build an Annotated Corpus + Guidelines for Annotation**

**Users Evaluation**

**Validation**

**SCAR Exporter + API**

**Data and Citation Contexts Extractor**

- bibliography entries
- section titles
- interest occurrence freq.

**Triple store**

- C1: Sections Classifier
  - Classifies sections according to DED ontology

**C2: Citation Functions Classifier**

- Uses supervised machine learning techniques, identifies citation functions
  - Uses method in
  - Uses data from
  - Extends
  - Cites as review

**XML**

- Full texts articles

**External data sources**

- Crossref
- ORCID
- CCO

**SCAR Exporter + API**

**External data sources**

- Crossref
- ORCID
- CCO

**Result of 3-fold cross validation using ground-truth dataset of the citation functions classifier.**

**Annotated Corpus**

- Annotated corpus about citation functions used to improve the performance of the classifier - 820 annotations (in progress)
- Ad-hoc guidelines in order to help annotators
- The annotated corpus will be soon released and freely available

**Would you like to participate?**

scar@cs.unibo.it