
OSCAR and LUCINDA with Wikidata

Ivan Heibi ivan.heibi2@unibo.it

Department of Classical Philology and Italian Studies, University of Bologna, Bologna, Italy

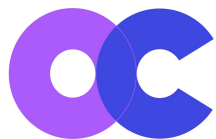
Silvio Peroni silvio.peroni@unibo.it

Department of Classical Philology and Italian Studies, University of Bologna, Bologna, Italy

David Shotton david.shotton@oerc.ox.ac.uk

Oxford e-Research Centre, University of Oxford, Oxford, UK

WikiCite 2018, Berkeley, California, USA



Who am I

● Bachelor and Master in Computer Science at the **University of Bologna, Italy**

● **Research fellow** under the supervision of **Silvio Peroni** on the “**OpenCitations** Enhancement Project”, funded by the Alfred P. Sloan Foundation

→ **Ph.D. student** at Department of Classical Philology and Italian Studies, University of Bologna, Bologna, Italy;
The application of Semantic Publishing technologies in the Science of Science research domain

OpenCitations

<http://opencitations.net/>

A scholarly infrastructure organization dedicated to **open scholarship** and the publication of open bibliographic and citation data by the use of **Semantic Web (Linked Data) technologies**, and engaged in advocacy for semantic publishing and **open citations**



OpenCitations Corpus (OCC)

~3 hundred thousand citing **bibliographic resources**
~14 million **citation links**
~7.5 million **cited resources**

COCI

~ 450 million **citation links**
~ 46 million **bibliographic resources**

Problem

No query interface that would permit **ordinary Web users (non experts in semantic web technologies)** to undertake queries and explore the datasets maintained by **OpenCitations**

Build a **Search** and **Browsing** interface to view the triplestore data, which performs SPARQL queries in its background and hides all the complexities of it to the users.



Reconfigure the developed tools to **Search** and **Browse** the **Wikidata RDF triplestore for scholarly data type of results**

OSCAR: the OpenCitations RDF Search Application

Firstly Presented at The SAVE-SD Workshop co-located with The Web Conference 2018

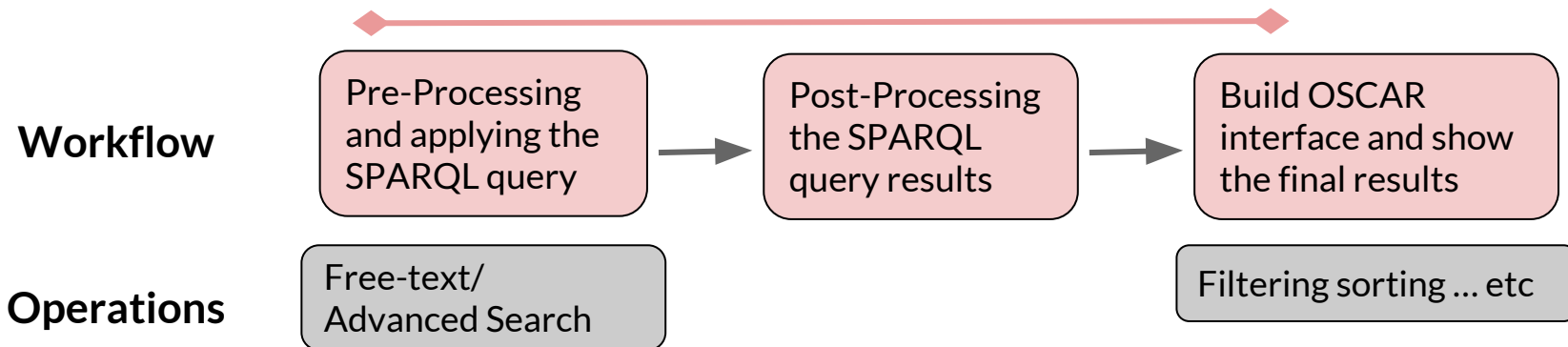
A **Free text search** (like the common Web search engines)

An **Advanced search** to perform complex queries logically connected

Filtering the result set retrieved

The interface, functionalities and queries must be **Customizable**

Configurable to work with other RDF triplestores, and to be **Integrated** as a new module inside a website



* For the larger detailed diagram [click here](#)

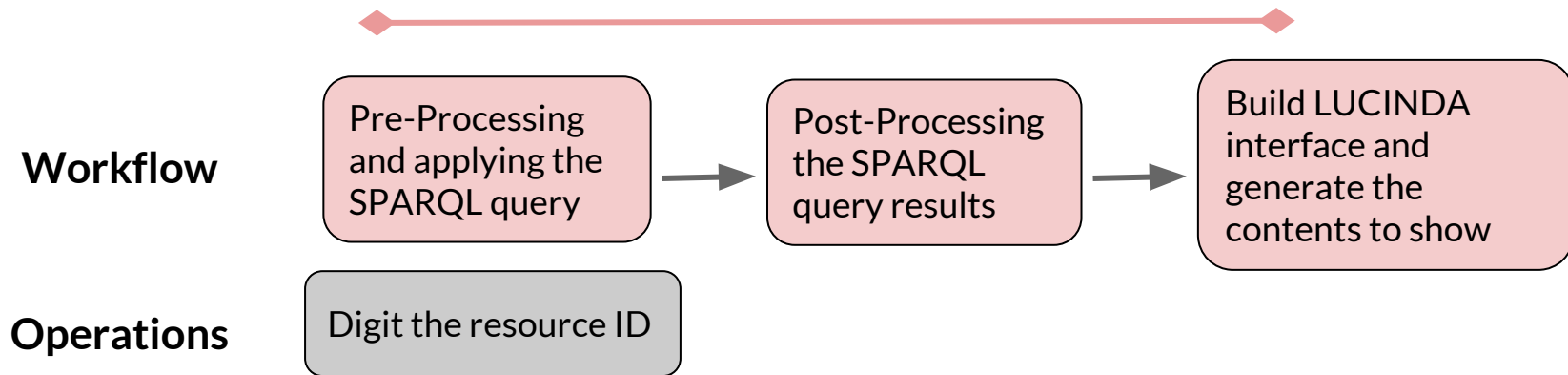
LUCINDA: the RDF data browser application

Browse the resource properties and related info by giving its **ID in the dataset**

Can **integrate OSCAR** inside it

The contents to show and queries to perform must be **Customizable**

Configurable to work with other RDF triplestores, and to be **Integrated** as a new module inside a website



WikiData: OSCAR and LUCINDA configuration

Querying only **scholarly articles (Q13442814)** related data

Two categories of results: **Articles (Q13442814)** and **Authors (P50)**

Using a **Free-text** and **Advanced** search

Get documents by: DOI, Q-ID, Journal ... etc.

Get authors by: ORCID, Q-ID, Occupation, Name, Article



A current demo of both OSCAR and LUCINDA with Wikidata is online and available at:

[OSCAR with Wikidata](https://tinyurl.com/yaqoogl2)

(<https://tinyurl.com/yaqoogl2>)

[LUCINDA with Wikidata](https://tinyurl.com/ybu8x5qb)

(<https://tinyurl.com/ybu8x5qb>)

* Check the configuration file of [OSCAR](#) and [LUCINDA](#) for Wikidata

How does OSCAR for Wikidata look like (1)

A screenshot from the demo available at the <https://tinyurl.com/yaqoogl2>

Scholarly article **Author**

10.1016/J.WEBSEM.2012.08.001

Which cite a specific article

And Or And Not

Remove

2002

With a minimum publication year

Add Rule +

Search the Wikidata Corpus

* Check OSCAR for Opencitations at <http://opencitations.net/search>

How does OSCAR for Wikidata look like (2)

A screenshot from the demo available at the <https://tinyurl.com/yaqoogl2>

Back to search

Number of rows per page: 10 Export results Sort: None

Limit to 17/17 results

< Fewer More >

All Show only Exclude

Select Authors ^

Select Date ^

Q-ID	Work title	Authors	Cited	Date
entity / Q24260641	Setting our bibliographic references free: towards open citation data	Silvio Peroni, Alexander Dutton, Tanya Gray, David Shotton	4	2015
entity / Q30098451	Describing Data Processing Pipelines in Scientific Publications for Big Data Injection	Sepideh Mesbah, Alessandro Bozzon, Christoph Lofi, Geert-Jan Houben	1	2017
entity / Q46412711	Automating semantic publishing	Silvio Peroni	0	2017
entity / Q46412887	Genuine semantic publishing	Tobias Kuhn, Michel Dumontier	0	2017

* Check OSCAR for Opencitations at <http://opencitations.net/search>

How does LUCINDA for Wikidata look like

A screenshot from the demo available at the <https://tinyurl.com/ybu8x5qb>.

Silvio Peroni

Author ORCID: [0000-0003-0530-4305](#)
Gender: male
Educated at: unknown
Occupation: computer scientist
Employer at: University of Bologna

Metrics

Author of **77** documents

Author's documents

Number of rows per page: [Export results](#) Sort:

Q-ID	Work title	Authors	Cited	Date
entity / Q24260683	FaBiO and CiTO: Ontologies for describing bibliographic resources and citations	Silvio Peroni, David Shotton	19	2012
entity / Q56459949	Modelling OWL Ontologies with Graffoo	Riccardo Falco, Aldo Gangemi, Silvio Peroni, David Shotton, Fabio Vitali	7	2014
entity / Q24260687	Scholarly publishing and linked data	Silvio Peroni, David Shotton, Fabio Vitali	6	2012

* Check LUCINDA for Opencitations at <http://opencitations.net/browser/ra/6932102>

Conclusions

We have presented **OSCAR**: the OpenCitations RDF Search Application, and
LUCINDA: the RDF data browser application.

Both are **configurable** tools for any triplestore providing a SPARQL endpoint url

We have built a customizable configuration of OSCAR and LUCINDA to query the
Wikidata triplestore



If you would like to use Oscar and lucinda please get in touch

Thanks for your attention

Ivan Heibi ivan.heibi2@unibo.it

Department of Classical Philology and Italian Studies, University of Bologna, Bologna, Italy

Silvio Peroni silvio.peroni@unibo.it

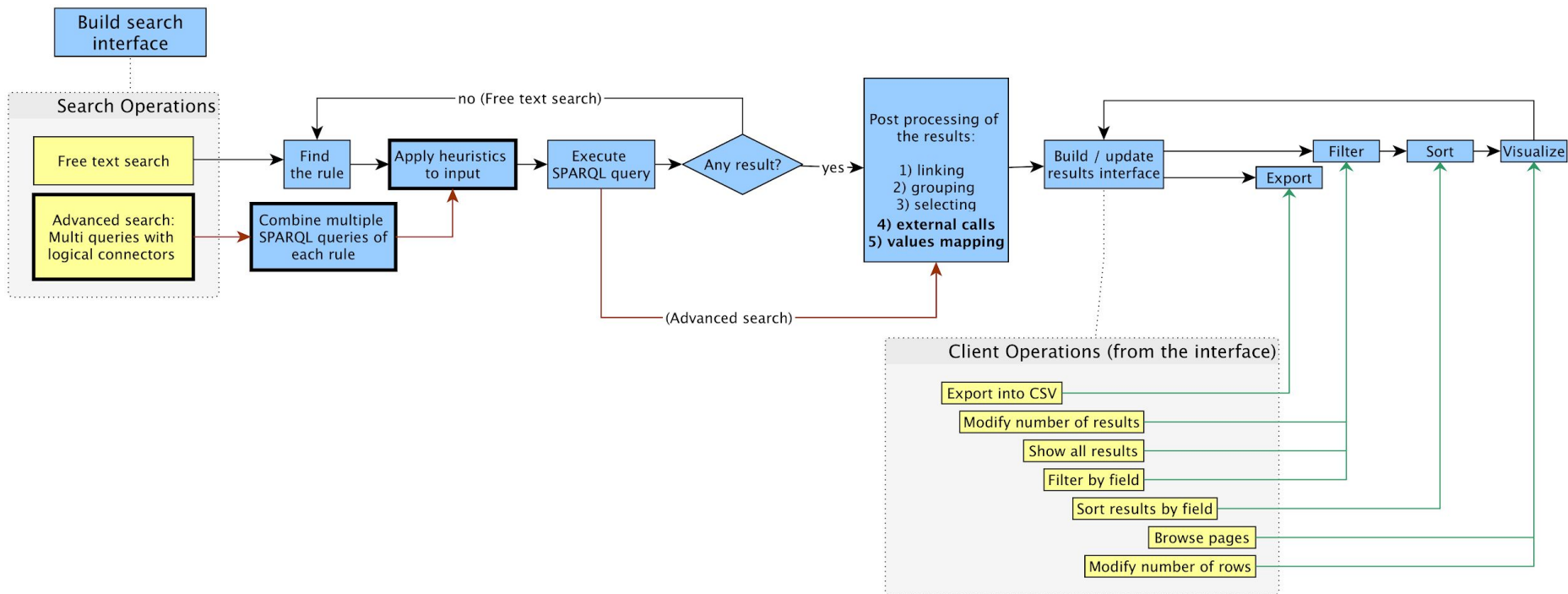
Department of Classical Philology and Italian Studies, University of Bologna, Bologna, Italy

David Shotton david.shotton@oerc.ox.ac.uk

Oxford e-Research Centre, University of Oxford, Oxford, UK

We gratefully acknowledge the financial support provided to us by the Alfred P. Sloan Foundation for the OpenCitations Enhancement Project (grant number G-2017-9800).

OSCAR workflow



OSCAR with WikiData: configuration file

```
"sparql_endpoint": "https://query.wikidata.org/sparql",

"prefixes": [
  {"prefix": "wd", "iri": "http://www.wikidata.org/entity/"},
  ...etc
],

"rules": [
  {
    "name": "qid",
    "category": "document",
    "regex": "Q\\d{1,}",
    "query":
      `{
        ?work wdt:P31 wd:Q13442814.
        BIND(<http://www.wikidata.org/entity/[[VAR]]> as ?work) .
      }`
    ...etc
  },
```

* Check the original entire file from [OSCAR repository](#)

LUCINDA with WikiData: configuration file

```
"sparql_endpoint": "https://query.wikidata.org/sparql",

"prefixes": [
  {"prefix":"wd","iri":"http://www.wikidata.org/entity/"},
  ...etc
],

"categories":{
  "document": {
    "rule": "Q.*",
    "query": ` SELECT DISTINCT ?work ?title ?doi ?pubmedid
              ?volume ?issue ?pages ?short_iri ?short_iri_id
              ?date (COUNT(distinct ?cites) AS ?out_cits)
              (COUNT(distinct ?cited) AS ?in_cits) ?author_resource
              ?author_short_iri ?author_str ?s_ordinal WHERE { ...etc
              `
  },
  ...etc
}
```

* Check the original entire file from [LUCINDA repository](#)