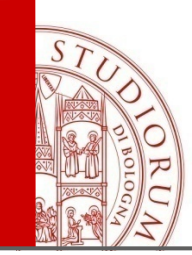


OSCAR: A customizable tool for free-text search over SPARQL endpoints

Ivan Heibi ivan.heibi2@unibo.it
Department of Computer Science and Engineering, 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

SAVE-SD 2018 - The Web Conference 2018 24 April 2018 - Lyon, France

Context: The RDF search application

- **RDF:** defines a data model in form of *subject-predicate-object* statements. **SPARQL:** the RDF query language.
- RDF datasets are successfully used and diffused for managing Linked Open Data datasets
- **However,** SPARQL is quite complex to learn and completely obscure to ordinary Web users.
- **Therefore,** provide a customizable search application interface, which performs SPARQL queries in background and hides all the complexities of it to the users.

OpenCitations

Creation and expansion of an open repository of scholarly citation data made available under a Creative Commons public domain dedication, which provides in RDF accurate citation information (bibliographic references) harvested from the scholarly literature.

- ~3 hundred thousand citing bibliographic resources
- ~12 million citation links
- ~6.5 million cited resources

➢ The repository is available for querying via its **SPARQL endpoint**

➢ It has **not had a query interface** that would permit ordinary Web users to undertake free text queries in order to explore the OC Corpus

OpenCitations web page

OSCAR

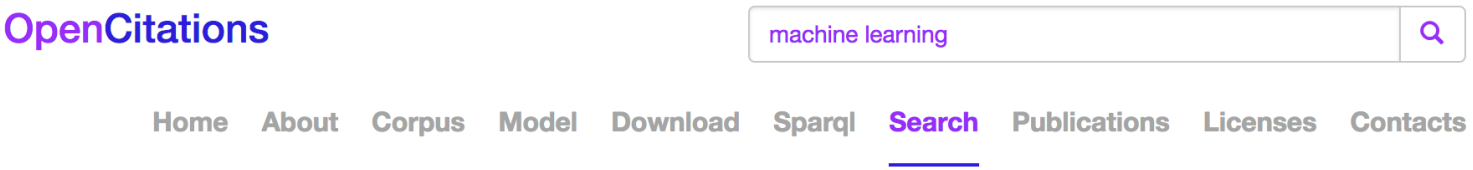
The *OpenCitations RDF Search Application:* a user-friendly search platform that can be used with any RDF triplestore providing a SPARQL endpoint. It's meant to work on the system client side, without the integration of external applications.

Requirements:

- Enable a **free text search**, such as is common to Web search engines
- Permit **filtering** of the result set retrieved
- The interface, functionalities and queries must be **customizable**
- Easily **configured** to work with other RDF triplestores (providing a SPARQL endpoint), and integrated as a new module inside a website

OSCAR is already available inside the OpenCitations website.
opencitations.net/search

OSCAR in OpenCitations



OSCAR in OpenCitations

OpenCitations

machine learning

Home About Corpus Model Download Sparql **Search** Publications Licenses Contacts

Number of rows per page: 5 Export results Sort: None

Limit to 155/311 results

All | Show only | Exclude

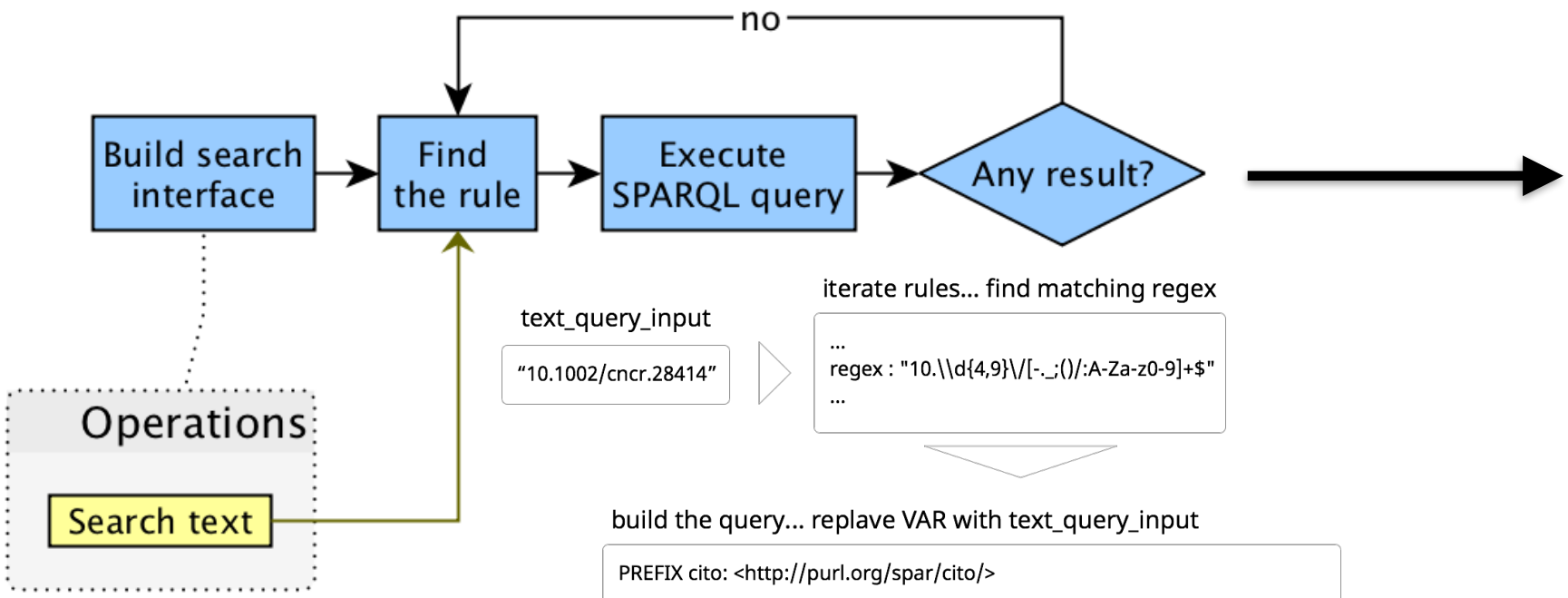
Select Year ∧

Select Authors ∧

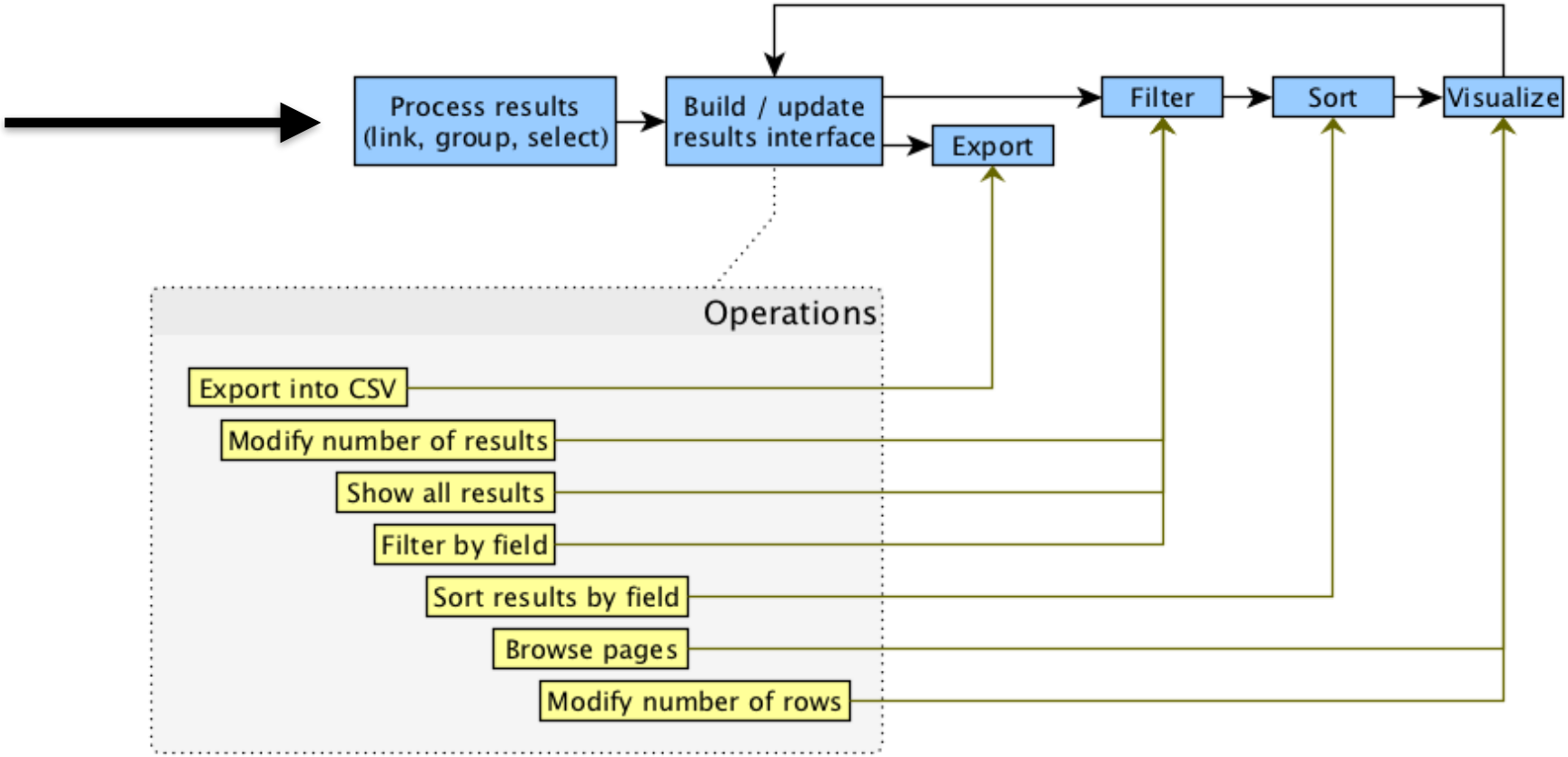
Corpus ID	Year	Title	Authors	Cited by
/br/6096958	2009	Machine Learning	Peter Flach	1
/br/40929		Machine Learning		0
/br/2713223		Machine Learning		0
/br/4231671	2008	Gaussian Processes for Machine Learning	Songthip T Ounpraseuth	1
/br/4453459	2014	DaDianNao: A Machine-Learning Supercomputer	Yunji Chen, Tao Luo, Shaoli Liu, Shijin Zhang, Liqiang He, Jia Wang, Ling Li, Tianshi Chen, Zhiwei Xu, Ninghui Sun, Olivier Temam	1

1 2 3 4 5 6 7 8 9 10 11 ... Next »

Workflow: SPARQL query selection



Workflow: results visualization



OSCAR usage

- To fully define OSCAR we use one particular **configuration file**
- **Use with different SPARQL endpoints:**
  We have defined three different OSCAR configuration files for three different projects: *OpenCitations Corpus,* **ScholarlyData** *and* **Wikidata**.
  - In all the three cases our aim was retrieving scholarly data type of results
  - The two main **Categories** retrieved were *Documents* and *Authors*
  - **Rules** to detect form a free-text input are: *DOIs, ORCIDs, and Title textual content.*

OSCAR usage: WikiData

Limit to 50/101 results

All | Show only | Exclude

Select Authors ∧

Select Date ∧

Number of rows per page: 5 Export results Sort: None

Resource IRI	Work title	Authors	Date
entity/Q27317312	On the universal structure of human lexical semantics.	Jon Wilkins	2016
entity/Q23712646	Wikidata as a semantic framework for the Gene Wiki initiative	Paul Pavlidis, Lynn Schriml, Sebastian Burgstaller-Muehlbacher, Elvira Mitraka, Timothy Elliott Putman, Andrew I. Su, Benjamin M. Good, Andra Waagmeester	2016
entity/Q24265390	Natural speech reveals the semantic maps that tile human cerebral cortex.		2016
entity/Q24289197	Neuroeconomic dissociation of semantic dementia and behavioural variant frontotemporal dementia		2016
entity/Q25894519	Exploring the role of the posterior middle temporal gyrus in semantic cognition: Integration of anterior temporal lobe with executive processes	Jonathan Smallwood	2016

1 2 3 4 5 6 7 8 9 10 Next »

OSCAR usage: ScholarlyData

Limit to 21/42 results

All | Show only | Exclude

Select Authors ∧

Number of rows per page: 5 Export results Sort: None

Resource IRI	Work title	Authors	DOI
eswc2010/paper/inuse/29	Generating innovation with semantically enabled TasLab portal	Pavel Shvaiko, Alessandro Oltramari, Davide Pozza, Giuseppe Angelini, Roberta Cuel	10.1007/978-3-642-13486-9_24
eswc2012/paper/demonstation/304	RDFaCE-Lite: a WYSIWYM editor for user-friendly semantic text authoring	Ali Khalili, Sören Auer, Soeren Auer	10.1007/978-3-662-46641-4_30
eswc2015/paper/in-use/111	Using semantic web technologies for enterprise architecture analysis	Bernhard Bauer, Maximilan Osenberg, Melanie Langermeier	10.1007/978-3-319-18818-8_41
www2012/demo/62	Round-trip semantics with Sztakipedia and DBpedia Spotlight	Pablo Mendes, Mihaly Heder, Mihály Héder	10.1145/2187980.2188048
www2012/demo/87	Automated semantic tagging of speech audio	Yves Raimond, Chris Lowis, Jonathan Tweed, Roderick Hodgson	10.1145/2187980.2188060

1 2 3 4 5 Next »

Empirical evaluation

User testing session to measure the usability, with 5 different users, testing OSCAR on the OpenCitations site

- *5 different tasks, and 2 different questionnaire: SUS (System Usability Scale) to measure the perceived system usability, open answers on usage experience.*

- **Outcomes:**
  Positive: SUS score was 87, large number of positive comments on the filtering operations provided.
  Negative: some ambiguity in some parts of OSCAR interface: on the filtering section, the default maximum number of results to visualize.

Conclusions

The outcomes notes have been all taken in consideration to improve OSCAR:

- New features and customizations
- A higher flexibility of the configuration file
- A visual re-adaption of the interface

Extensions and related projects:

- Complex combinations of field-oriented queries by means of logical connections (*OR*, *AND*, etc.)
- Developing a related tool, named *LUCINDA*, to browse the triplestore resources

Thanks for your attention

Use OSCAR !
And help us improve it.

Ivan Heibi ivan.heibi2@unibo.it
Department of Computer Science and Engineering, 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).

Thanks for your attention

Use OSCAR (not Corcho) !
And help us improve it.

Ivan Heibi ivan.heibi2@unibo.it
Department of Computer Science and Engineering, 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).