Open
Research
Knowledge
Graph

What is ORKG?

Yaser Jaradeh @ L3S/TIB

DaMaLOS 2020 Workshop

A view on scholarly communication

1620

1883

2020

ORKG

A view on scholarly communication

1620   1883   2020

What about Metadata?

Well.. some things changed

(MAG, Crossref, Wikidata, WikiCite, Researchgate, Semantic Scholar)

ORCID

PDF

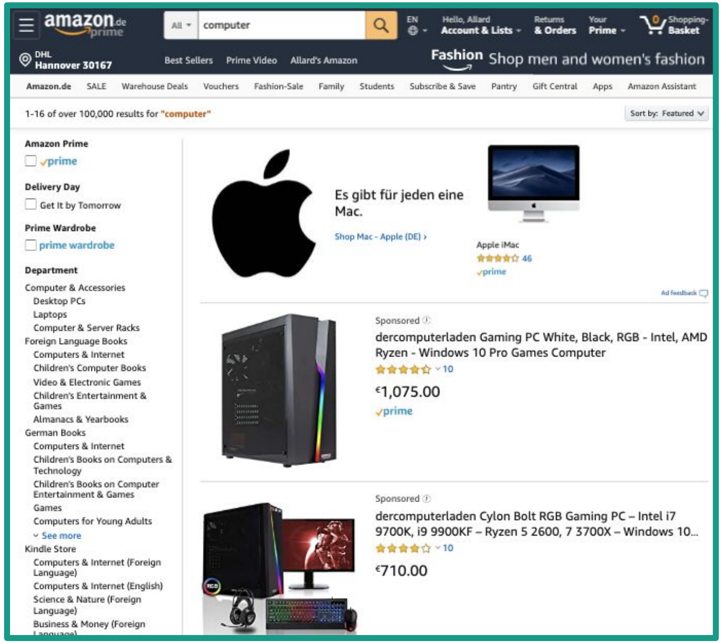
But what about the content ?

(i.e., the contributions)

Let’s first think about it (what about other domains?)

Let’s first think about it (what about other domains?)

Let's first think about it (what about other domains?)



Hmm interesting (still other domains?)

Hmm interesting (still other domains?)

Hmm interesting (still other domains?)

They didn’t just made a digital format, they revolutionized the domain

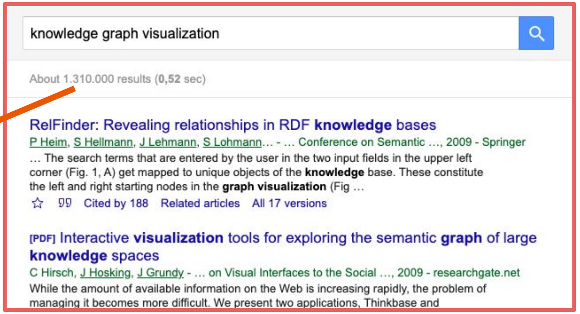
What about scholarly domain?

Finding and comparing scientific literature is time consuming

What about scholarly domain?

Finding and comparing scientific literature is time consuming

Finding the right papers



1.310.000 results

Comparing results

"Our visualization tool is able to visualize graphs, charts and trees"

"The focus lies on visualizing graphs"

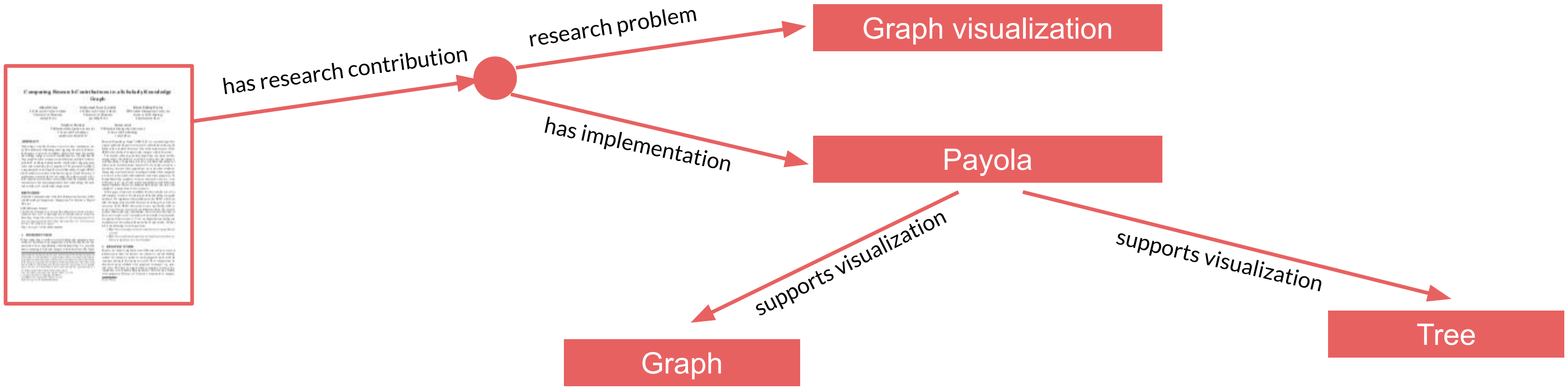
The solution

We need structured research data

The Open Research Knowledge Graph (ORKG) focuses on making research papers structured by using mainly crowdsourcing

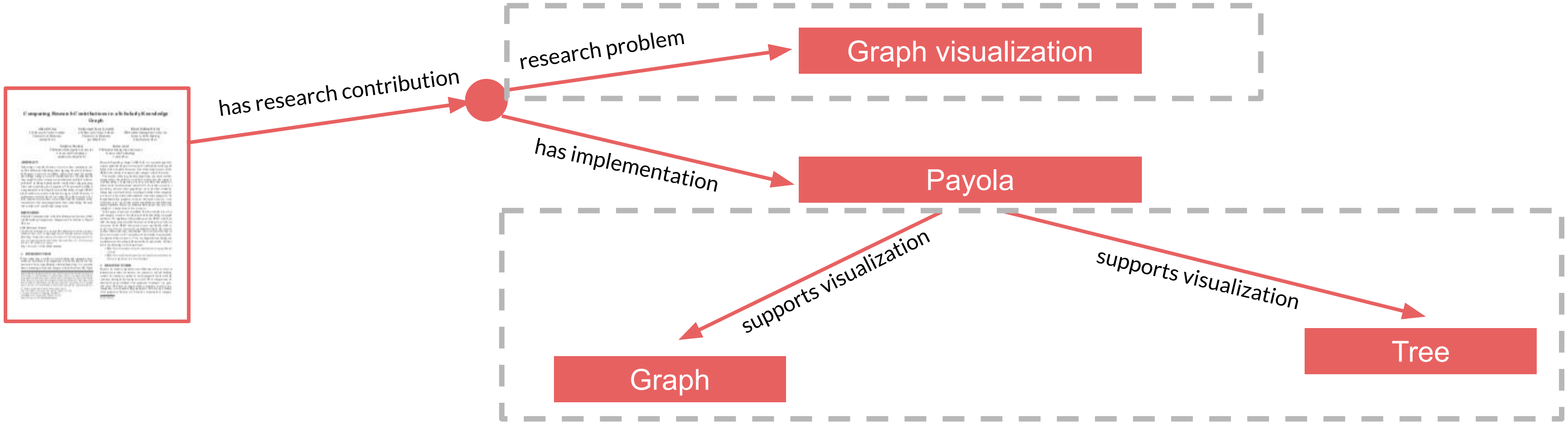
Open Research Knowledge Graph

- The **Open Research Knowledge Graph (ORKG)** focuses on making research papers structured by using mainly **crowdsourcing**



Open Research Knowledge Graph

- The **Open Research Knowledge Graph (ORKG)** focuses on making research papers structured by using mainly **crowdsourcing**



Crowdsourcing, really?

- For generating the structured paper descriptions
- Why: automated methods (e.g., NLP) are not accurate enough to generate a high quality knowledge graph
- Who: paper authors and domain experts

Why ORKG?


IEEE TRANSACTIONS ON INTELLIGENT TRANSPORTATION SYSTEMS, VOL. 15, NO. 4, AUGUST 2014 1441

Situational Knowledge Representation for Traffic Observed by a Pavement Vibration Sensor Network

Markus Stocker, Mauno Rönkkö, and Mikko Kolehmainen

***Abstract*—Information systems that build on sensor networks often process data produced by measuring physical properties. These data can serve in the acquisition of knowledge for real-world situations that are of interest to information services and, ultimately, to people. Such systems face a common challenge, namely the considerable gap between the data produced by measurement and the abstract terminology used to describe real-world situations. We present and discuss the architecture of a software system that utilizes sensor data, digital signal processing, machine learning, and knowledge representation and reasoning to acquire, represent, and infer knowledge about real-world situations observable by a sensor network. We demonstrate the application of the system to vehicle detection and classification by measurement of road pavement vibration. Thus, real-world situations involve vehicles and information for their type, speed, and driving direction.**

***Index Terms*—Knowledge acquisition, knowledge representation, machine learning, sensor data, sensor networks, traffic monitoring.**

I. Introduction

We propose a software system architecture and implementation for the continuous and automated representation of knowledge for real-world situations observable by a sensor network. In this paper, we demonstrate the application of the software system to intelligent transportation systems. Thus, real-world situations involve vehicles and information for their type, speed, and driving direction.

According to Finkelstein [1], "measurement is the process of empirical, objective, assignment of numbers to properties of objects or events of the real world in such a way as to describe them." A sensor is a device that performs measurement, in that it transforms the signal of a physical property (e.g., heat) into numbers or, more generally, into data [2]. Sensor measurement is, hence, the process of recurrent application of such transformation for certain temporal and spatial locations. The result of sensor measurement is sensor data. Sensor data represent the change of the signal over time.

Despite recent advancements in sensor data management, processing, and query [2]–[4], as well as semantic description of sensors and data [5]–[7], making sense of sensor data is an ongoing challenge [8]–[10] because of the difference in the degree to which sensor data represents information about a signal and information about, or related to, a physical property [11]. In other words, it is a challenge because of the considerable gap between data produced by measurement and abstract terminology [12] used by people to describe (the properties of) real-world objects or events.

We are interested in *situations* involving real-world objects that affect a physical property, for which a signal is measured by means of sensors. In this paper, vehicles are the real-world objects and road pavement vibration is the physical property. We present the architecture of a software system that utilizes digital signal processing, machine learning, and knowledge representation and reasoning to acquire, represent, and infer knowledge about real-world situations involving vehicles. The system aims at reducing the gap between road pavement vibration measurement data and abstract terminology used to describe real-world situations involving vehicles.

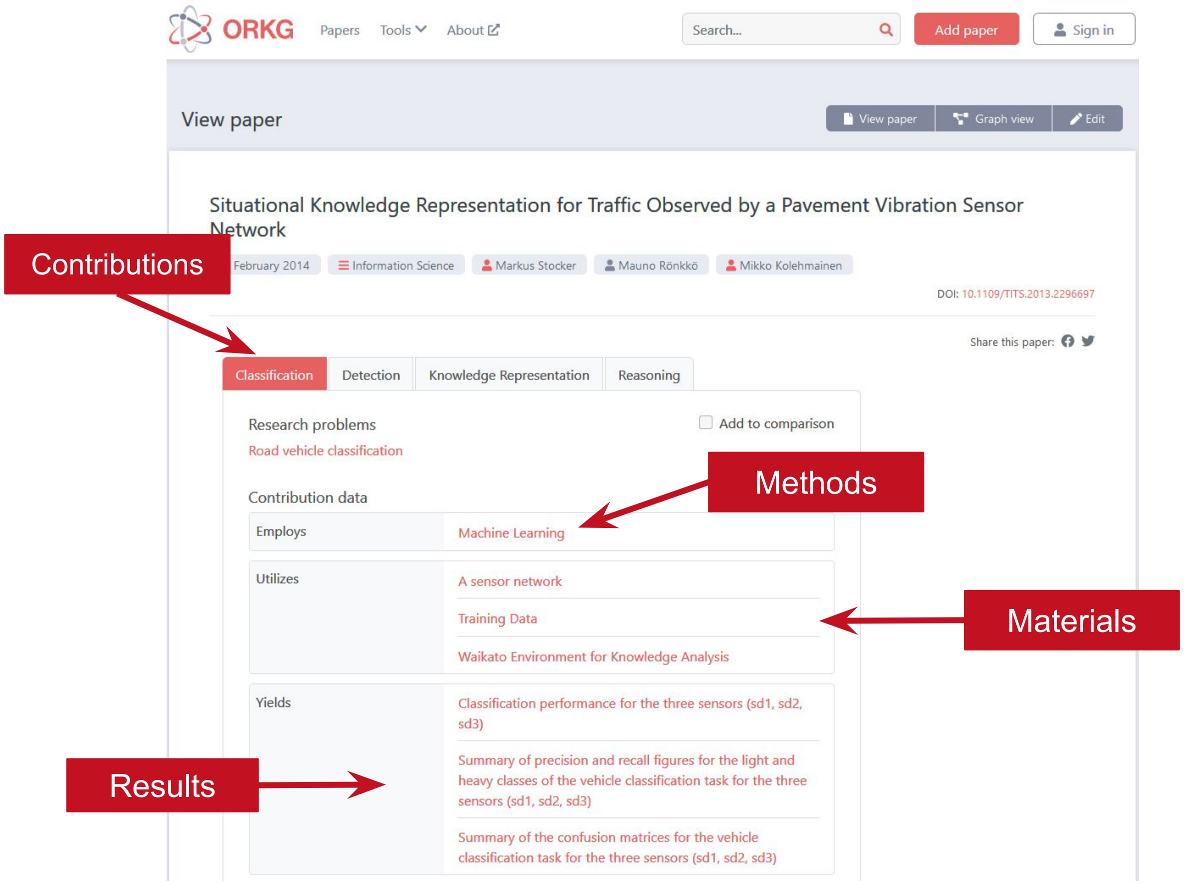
Digital signal processing techniques are iteratively applied to a sliding window over sensor data to enhance the vibration signal and to transform sensor data (time domain) into patterns (frequency domain). Machine learning is used to classify patterns. We employ multilayer perceptron (MLP) feedforward artificial neural networks [13]. Techniques in knowledge representation are utilized to formally represent domain concepts, instances, and relations. A concept of interest to our domain is the vibration sensor. The (installed) sensors are represented as instances of this concept. An instance may have a number of relations, e.g., to a spatial location. We represent sensors and observations using the Semantic Sensor Network Ontology (SSNO) [14].[1] SSNO is an "ontology for describing the capabilities of sensors, the act of sensing and the resulting observations" [15]. We employ the Situation Theory Ontology[2] (STO) [16] to represent knowledge about real-world situations, which are acquired from observations. The STO captures the key aspects of the situation theory developed by Barwise and Perry [17] and extended by Devlin [18]. The theory relates to the work on situation awareness by Endsley [19], [20] as it encompasses most of the concepts discussed in [16]. Both the SSNO and the STO serve as upper ontologies from which we extend to accommodate domain knowledge. The hybrid use of the SSNO and the STO allows for a multilevel abstraction of sensor measurement data and the use of appropriate terminology and formalization at each level.

Manuscript received April 12, 2013; revised August 16, 2013 and November 20, 2013; accepted December 22, 2013. Date of publication February 4, 2014; date of current version August 1, 2014. The infrastructure to access and collect vibration and camera data, as well as the data, are part of research funded by Tekes, the Finnish Funding Agency for Technology and Innovation (funding decision number 40075/09). The Associate Editor for this paper was P. Grisleri.

The authors are with the Department of Environmental Science, University of Eastern Finland, 70211 Kuopio, Finland (e-mail: markus.stocker@uef.fi; mauno.ronkko@uef.fi; mikko.kolehmainen@uef.fi).

Digital Object Identifier 10.1109/TITS.2013.2296697

[1] http://purl.oclc.org/NET/ssnx/ssn
[2] http://vistology.com/ont/2008/STO/STO.owl



Why ORKG?

Contribution comparison 3

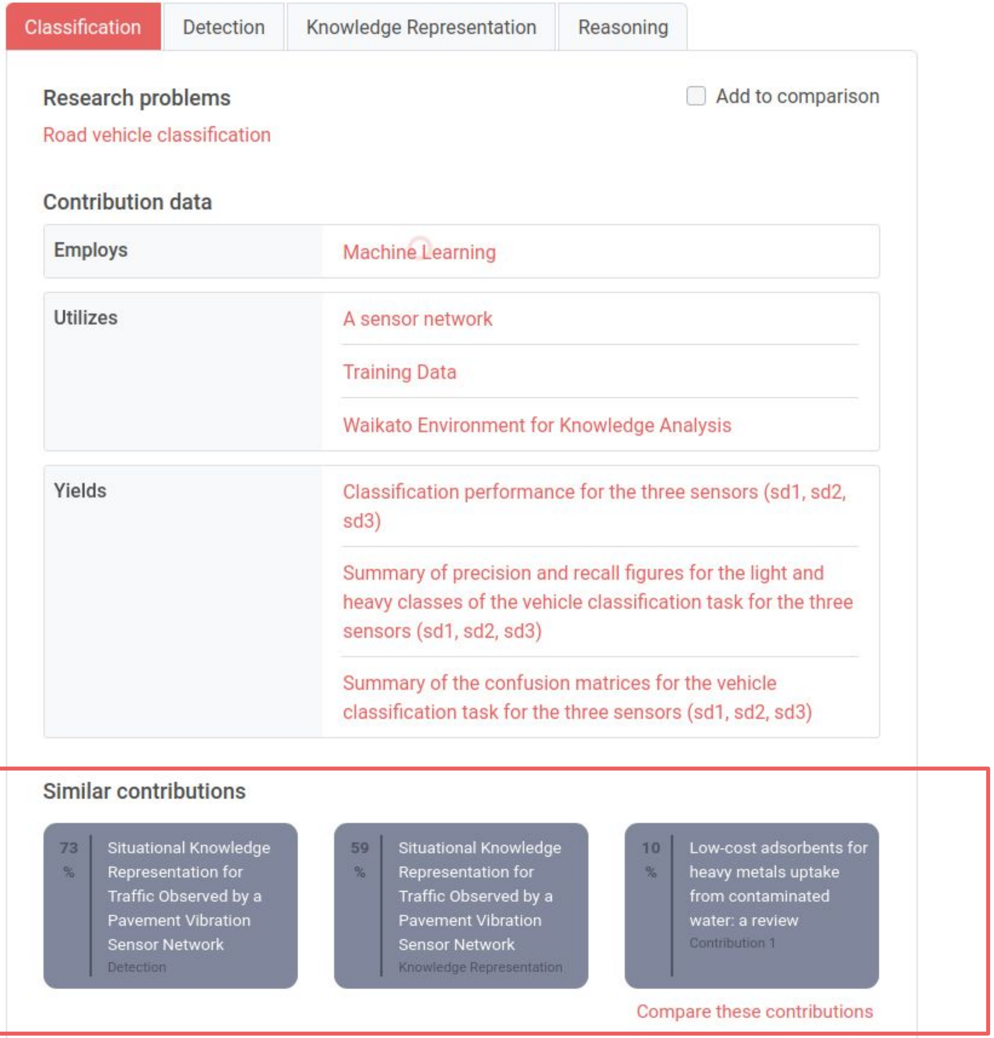
View | Add contribution | More

Covid-19 Reproductive Number Estimates

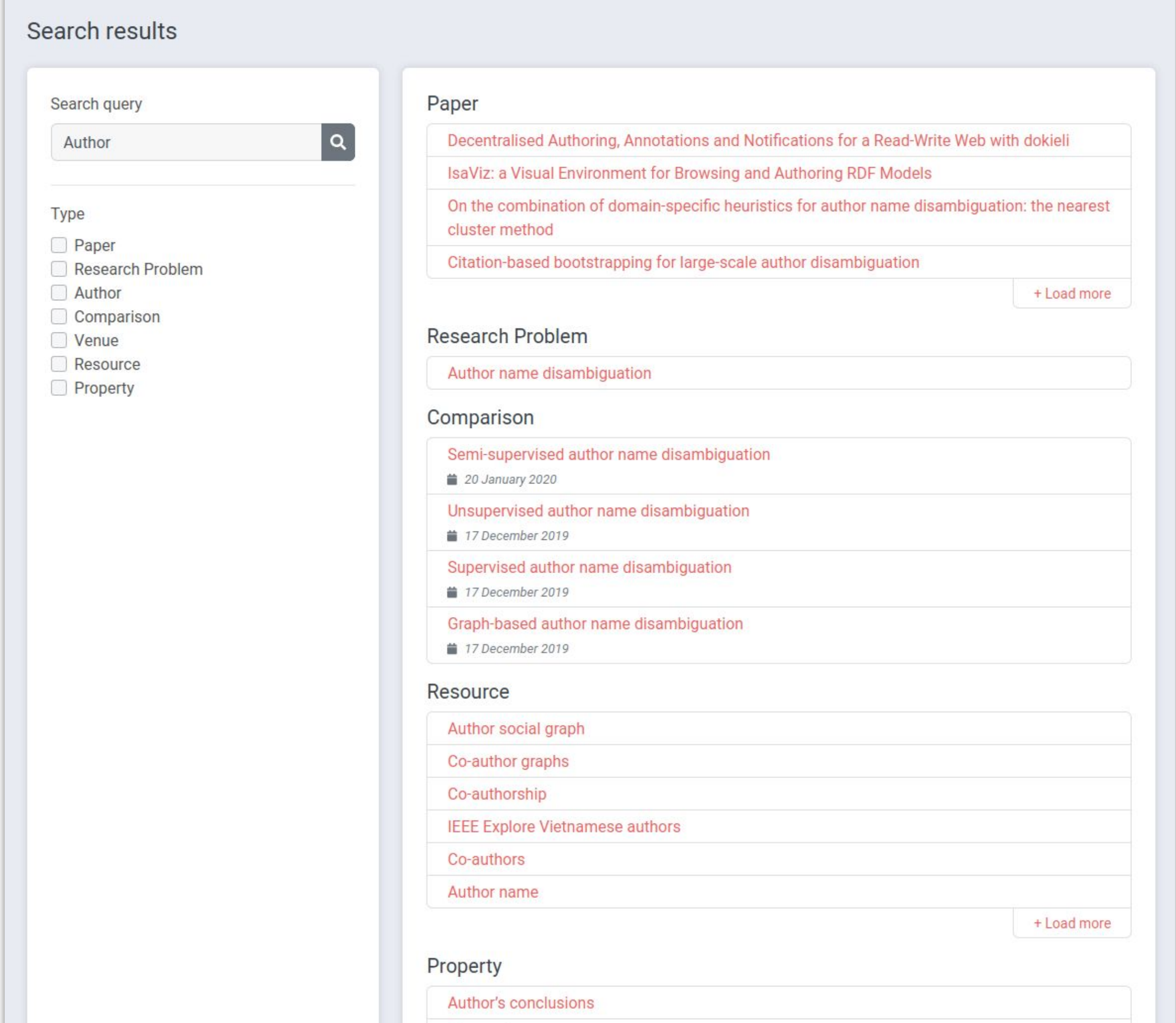
Method: Intelligent merge

Properties	The early phase of the COVID-19 outbreak in Lombardy, Italy Contribution 1 - 2020	Transmission potential of COVID-19 in Iran Contribution 1 - 2020	Transmission potential of COVID-19 in Iran Contribution 2 - 2020
Has value	3.1	3.6	3.58
Location	Lombardy, Italy	Iran	Iran
Confidence interval (95%)	Confidence interval (95%)	Confidence interval (95%)	Confidence interval (95%)
Lower confidence limit	2.9	3.4	1.29
Upper confidence limit	3.2	4.2	8.46
Has beginning	2020-01-14	2020-02-19	2020-02-19
Has end	2020-03-08	2020-02-29	2020-02-29

Want another reason?



Give me more reasons!



Wow! also this

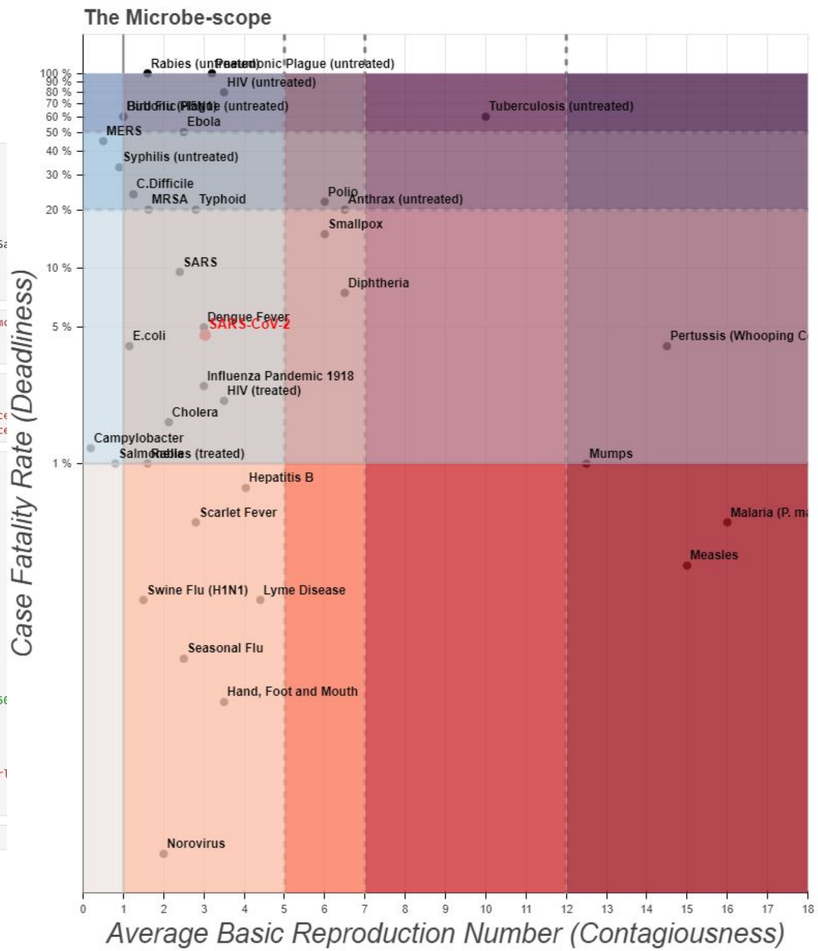
```
import requests
import datetime
import pandas as pd
import numpy as np
from orkg import ORKG
from bokeh.io import export_png
from bokeh.models import ColumnDataSource, HoverTool, WheelZoomTool, ResetTool, Sa
from bokeh.plotting import figure, show, output_notebook

output_notebook()
```

```
orkg = ORKG(host='https://orkg.org/orkg', simcomp_host='https://orkg.org/orkg/simc

df = orkg.contributions.compare_dataframe(comparison_id='R44930')
```

```
dates = np.array([datetime.date.fromisoformat(x) for x in df.loc['has end', :]])
values = np.float32(df.loc['Has value', :])
lower = np.array([np.float32(x) if x else np.nan for x in df.loc['Lower confidence
upper = np.array([np.float32(x) if x else np.nan for x in df.loc['Upper confidence
```

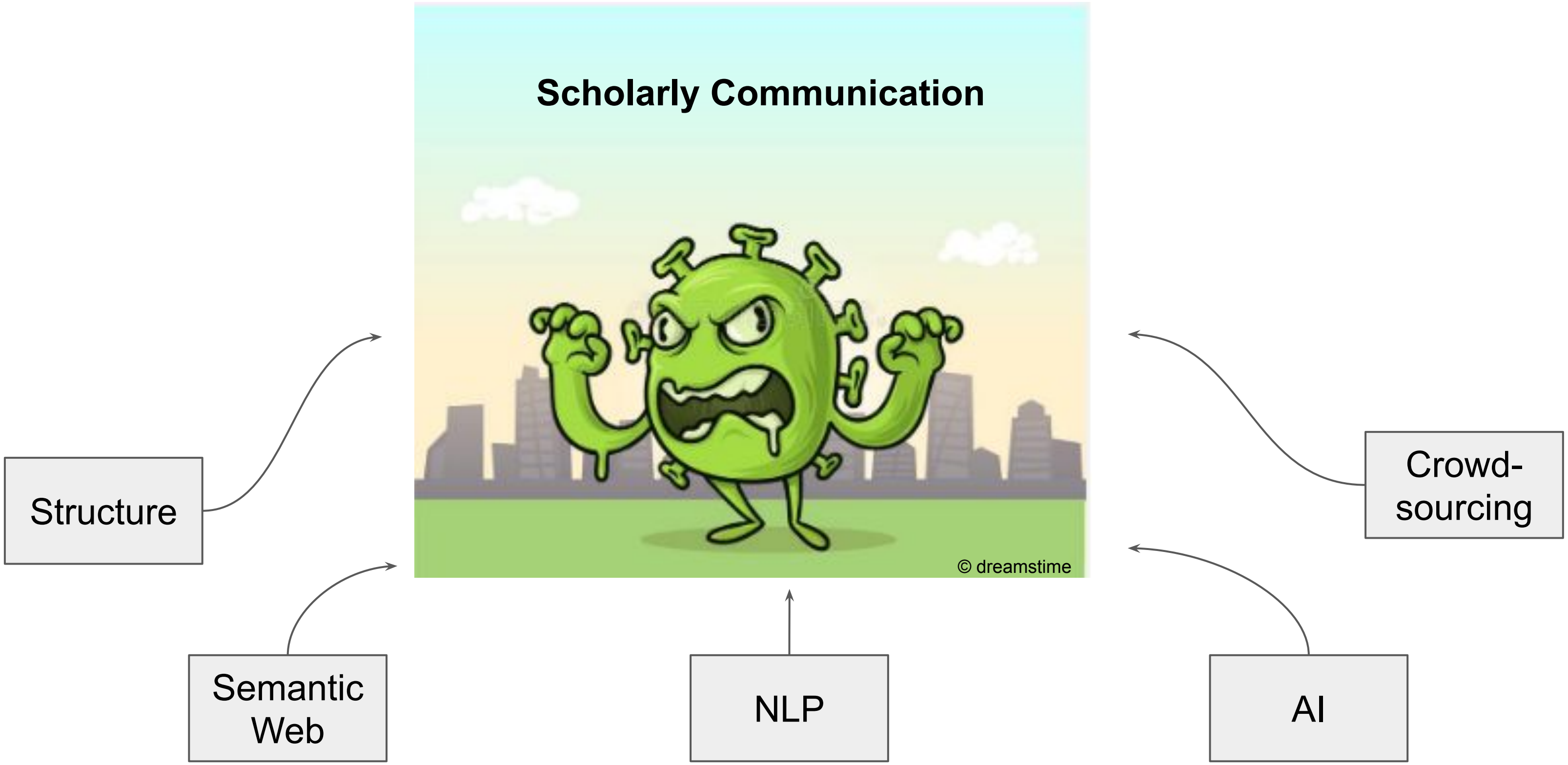
```
hover1 = HoverTool(
    tooltips=[
        ('Date', '@date{%F}'),
        ('R0', '@value{0.ff}'),
        ('95% CI', '@lower{0.ff}-@upper{0.ff}')
    ],
    formatters={
        '@date': 'datetime',
        '@{value}' : 'printf',
        '@{lower}' : 'printf',
        '@{upper}' : 'printf'
    }
)

df = pd.DataFrame(data=dict(date=dates, value=values, lower=lower, upper=upper))
source = ColumnDataSource(df)
p = figure(x_axis_type="datetime", y_range=(0, 9), plot_width=800, plot_height=35(
p.xaxis.formatter=DatetimeTickFormatter(days=['%d %b'])
p.yaxis.axis_label = 'basic reproduction number'
p.circle('date', 'value', source=source, size=7, color='purple')
p.add_layout(
    Whisker(source=source, base='date', upper='upper', lower='lower', level='overl
)
show(p)
```

```
export_png(p, filename='img/R0-estimates-plot.png')
```

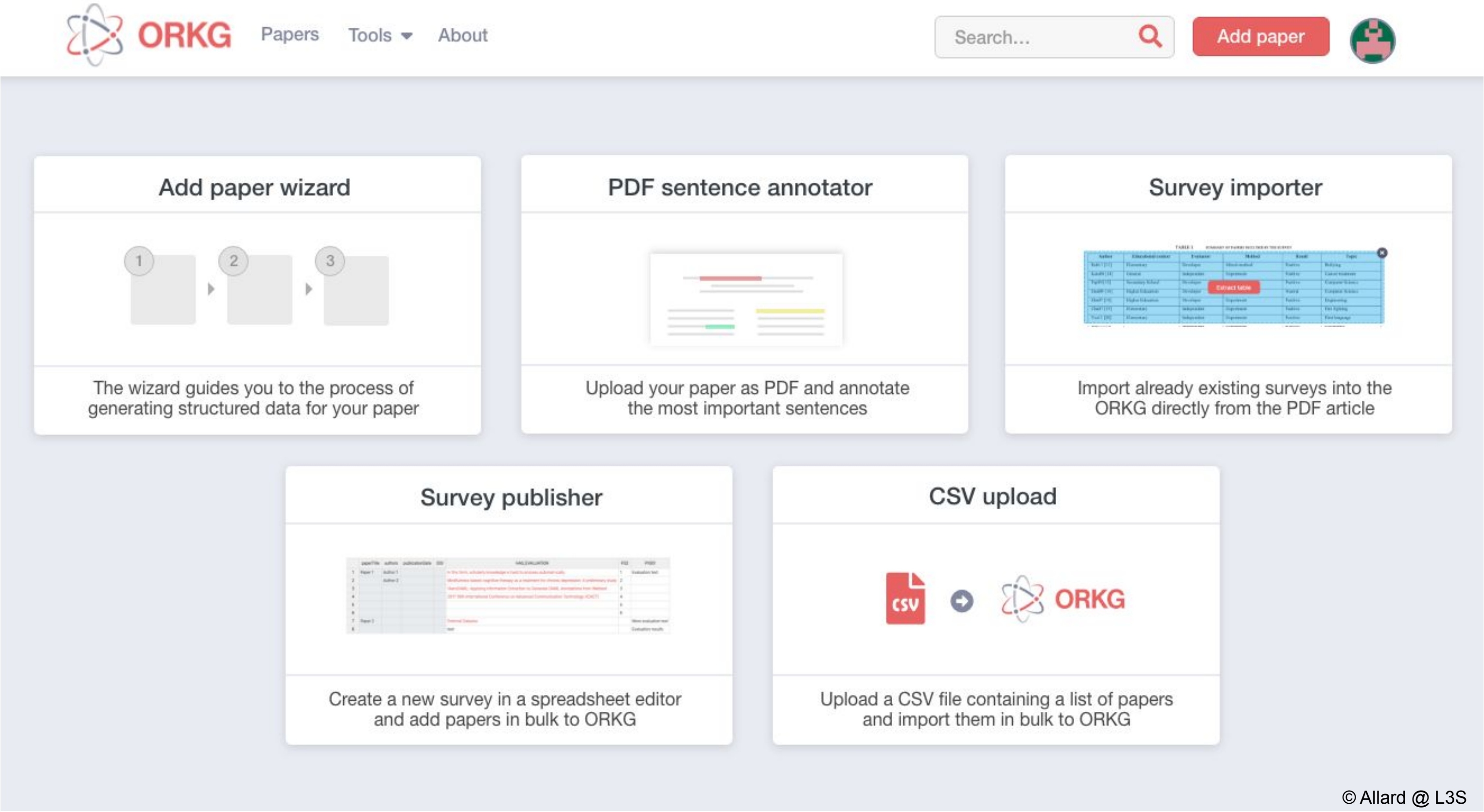


cf. https://gitlab.com/TIBHannover/orkg/orkg-notebooks

How to add papers into ORKG

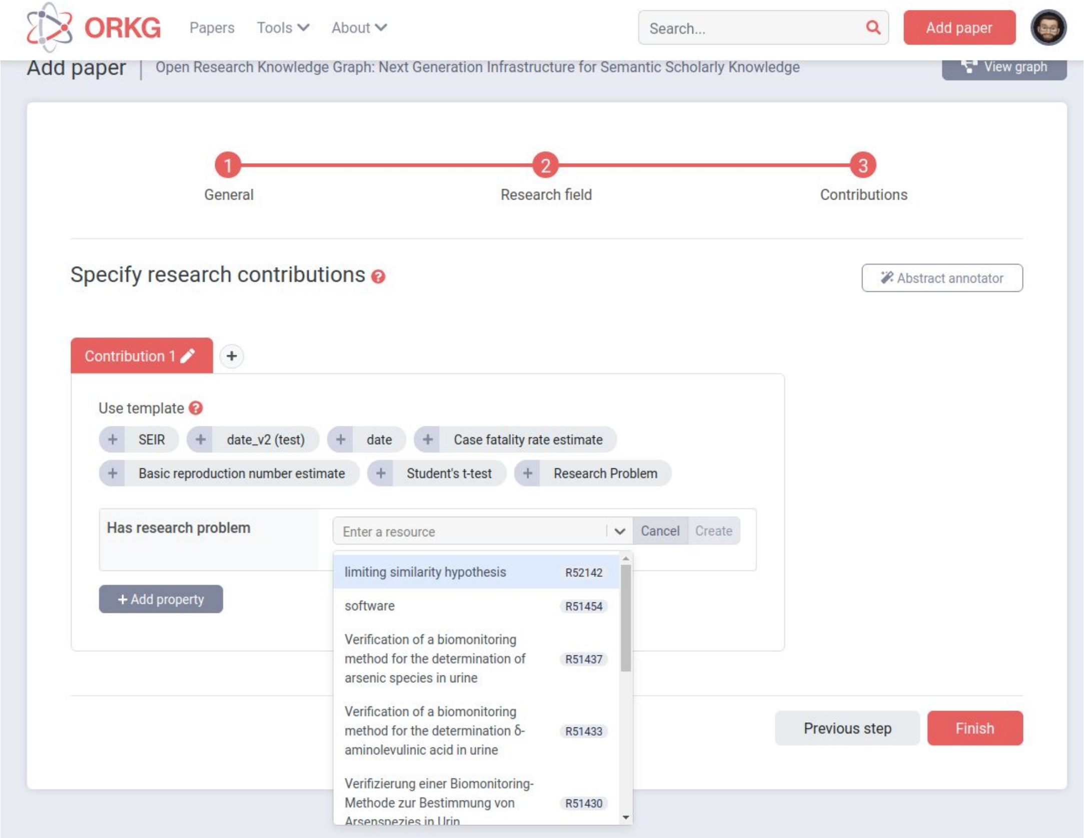


How to add papers into ORKG

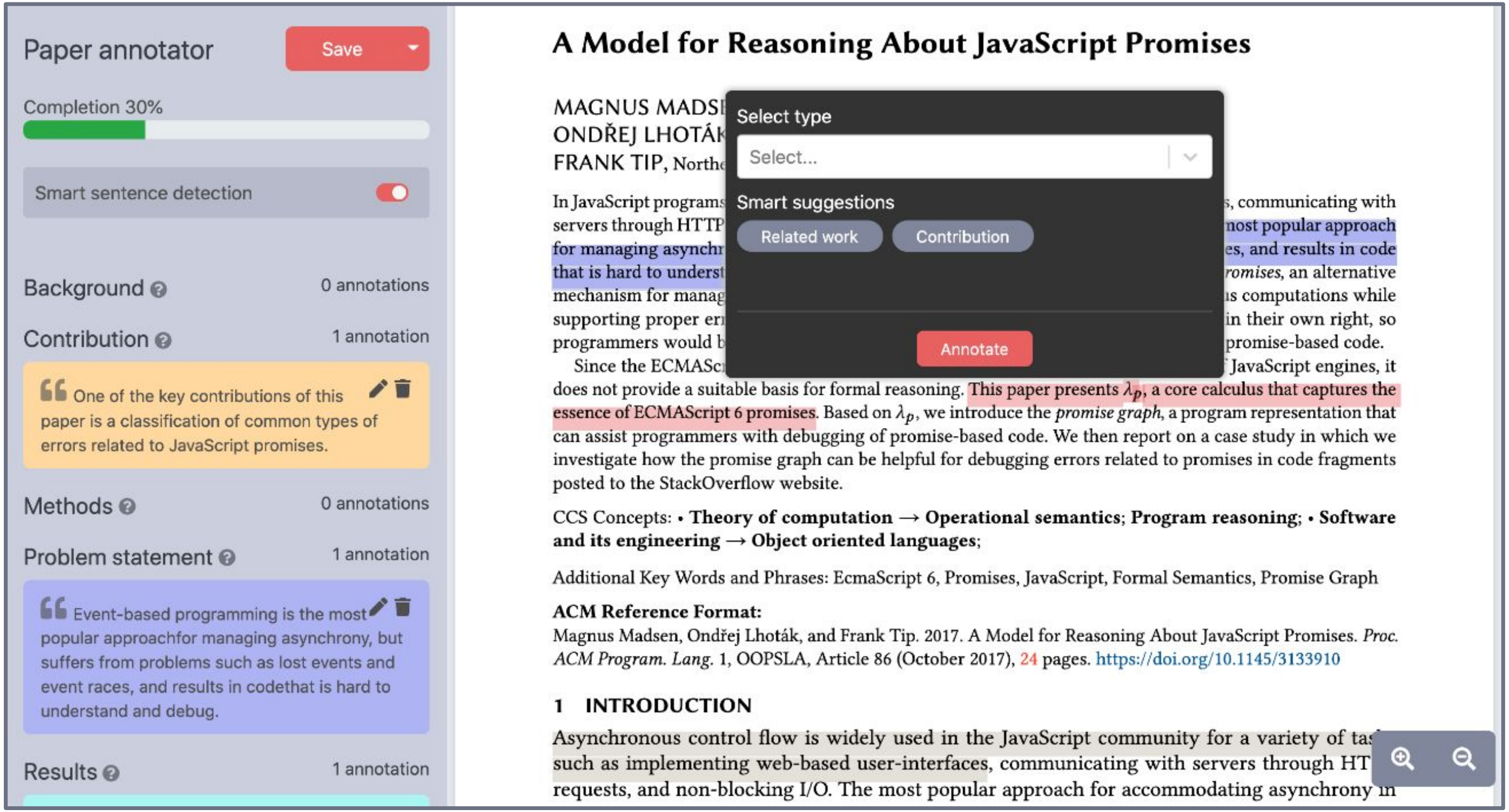


© Allard @ L3S

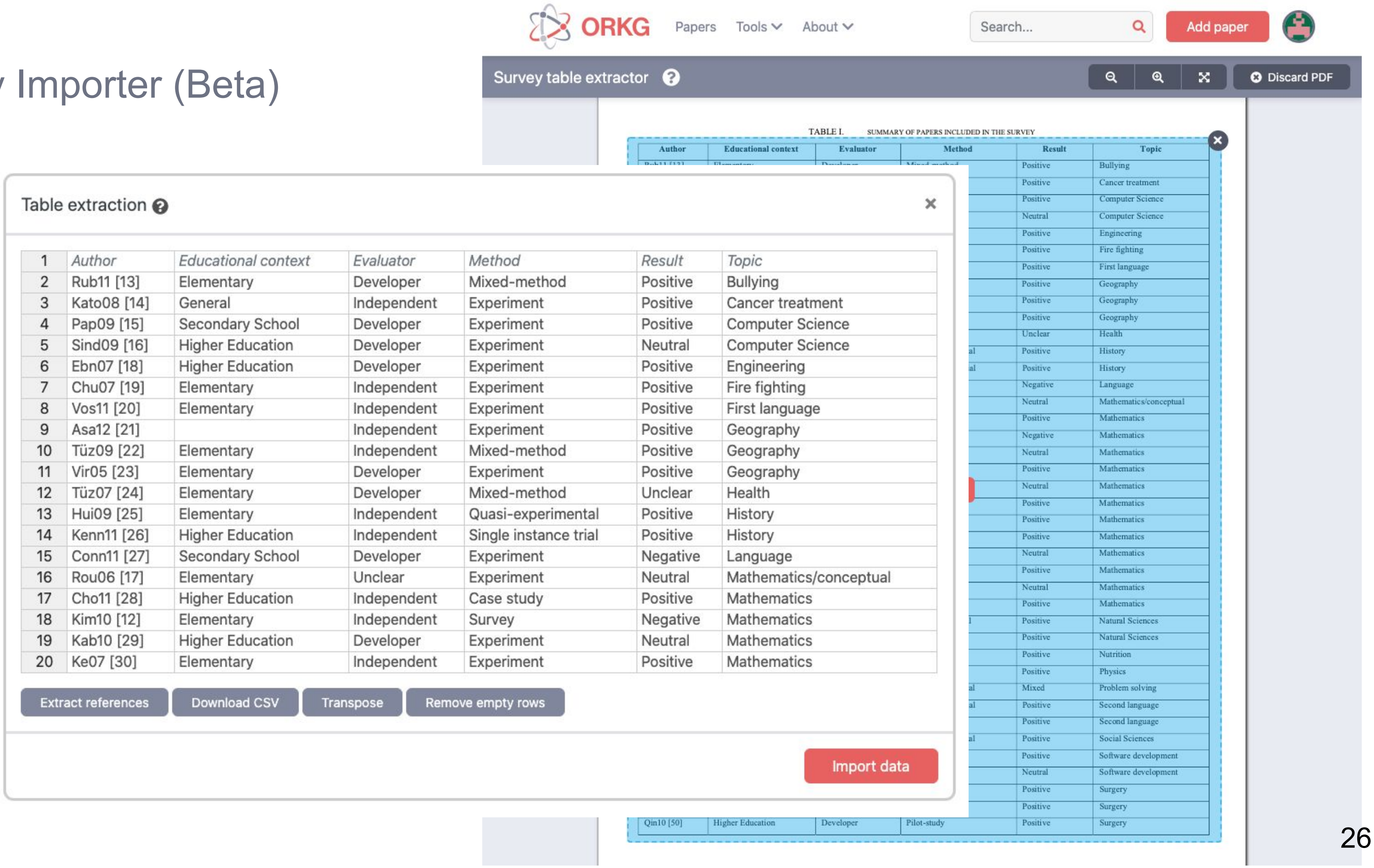
Paper Wizard



PDF Annotator (Alpha)



Survey Importer (Beta)



1	*Author*	*Educational context*	*Evaluator*	*Method*	*Result*	*Topic*
2	Rub11 [13]	Elementary	Developer	Mixed-method	Positive	Bullying
3	Kato08 [14]	General	Independent	Experiment	Positive	Cancer treatment
4	Pap09 [15]	Secondary School	Developer	Experiment	Positive	Computer Science
5	Sind09 [16]	Higher Education	Developer	Experiment	Neutral	Computer Science
6	Ebn07 [18]	Higher Education	Developer	Experiment	Positive	Engineering
7	Chu07 [19]	Elementary	Independent	Experiment	Positive	Fire fighting
8	Vos11 [20]	Elementary	Independent	Experiment	Positive	First language
9	Asa12 [21]		Independent	Experiment	Positive	Geography
10	Tüz09 [22]	Elementary	Independent	Mixed-method	Positive	Geography
11	Vir05 [23]	Elementary	Developer	Experiment	Positive	Geography
12	Tüz07 [24]	Elementary	Developer	Mixed-method	Unclear	Health
13	Hui09 [25]	Elementary	Independent	Quasi-experimental	Positive	History
14	Kenn11 [26]	Higher Education	Independent	Single instance trial	Positive	History
15	Conn11 [27]	Secondary School	Developer	Experiment	Negative	Language
16	Rou06 [17]	Elementary	Unclear	Experiment	Neutral	Mathematics/conceptual
17	Cho11 [28]	Higher Education	Independent	Case study	Positive	Mathematics
18	Kim10 [12]	Elementary	Independent	Survey	Negative	Mathematics
19	Kab10 [29]	Higher Education	Developer	Experiment	Neutral	Mathematics
20	Ke07 [30]	Elementary	Independent	Experiment	Positive	Mathematics

Survey Builder (WIP)

Survey builder | View survey | Finish

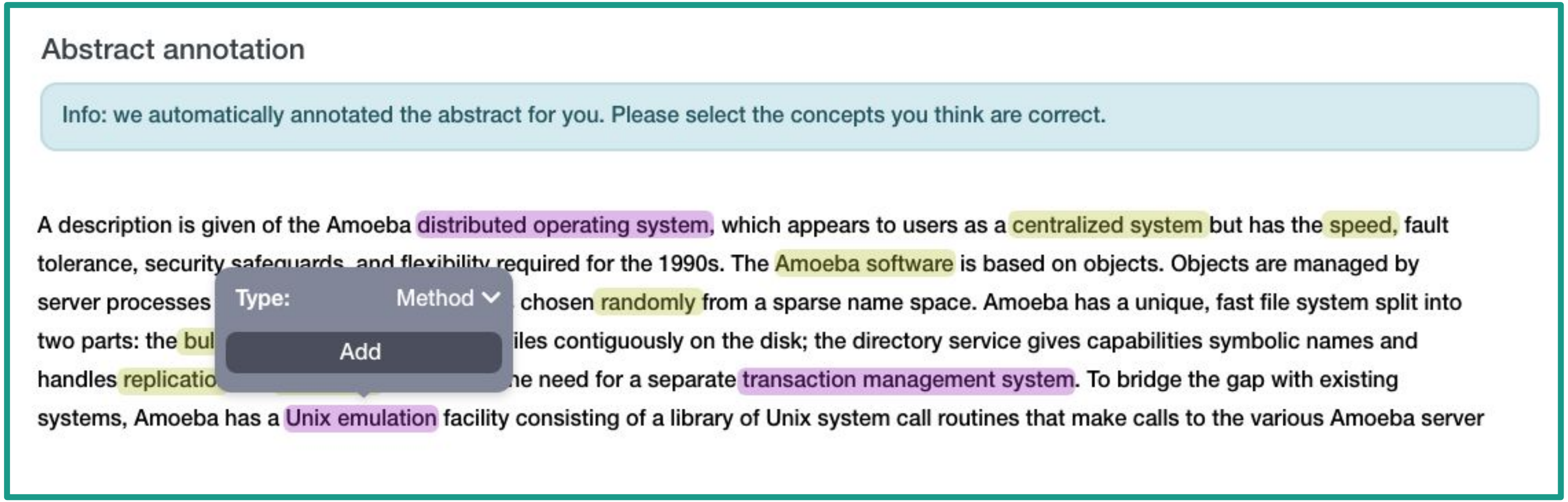
Insert row above | Insert row below | Remove row | Insert column left | Insert column right | Remove column

	Title	Authors	Publication date	DOI	Algorithm	Population
1	Algorithm and hardware for a merge sort using	S. Todd	2019-06	10.1016/2017.11.022	Merge sort	234,543,457
2	A variant of heapsort with almost optimal…	X. Svante	2004-06	10.2336/203.43.2	Heap sort	256,565,655
		C. Carlsson				242,333,343
		H. Persons				345,444,566
3	Bubble sort: an archaeologic alalgorithmic	Owen Astrachan	1997-06	10.5445/23.33.3332	Bubble sort	878,434,111
		H. Secondsson				122,344,222
						211,886,664

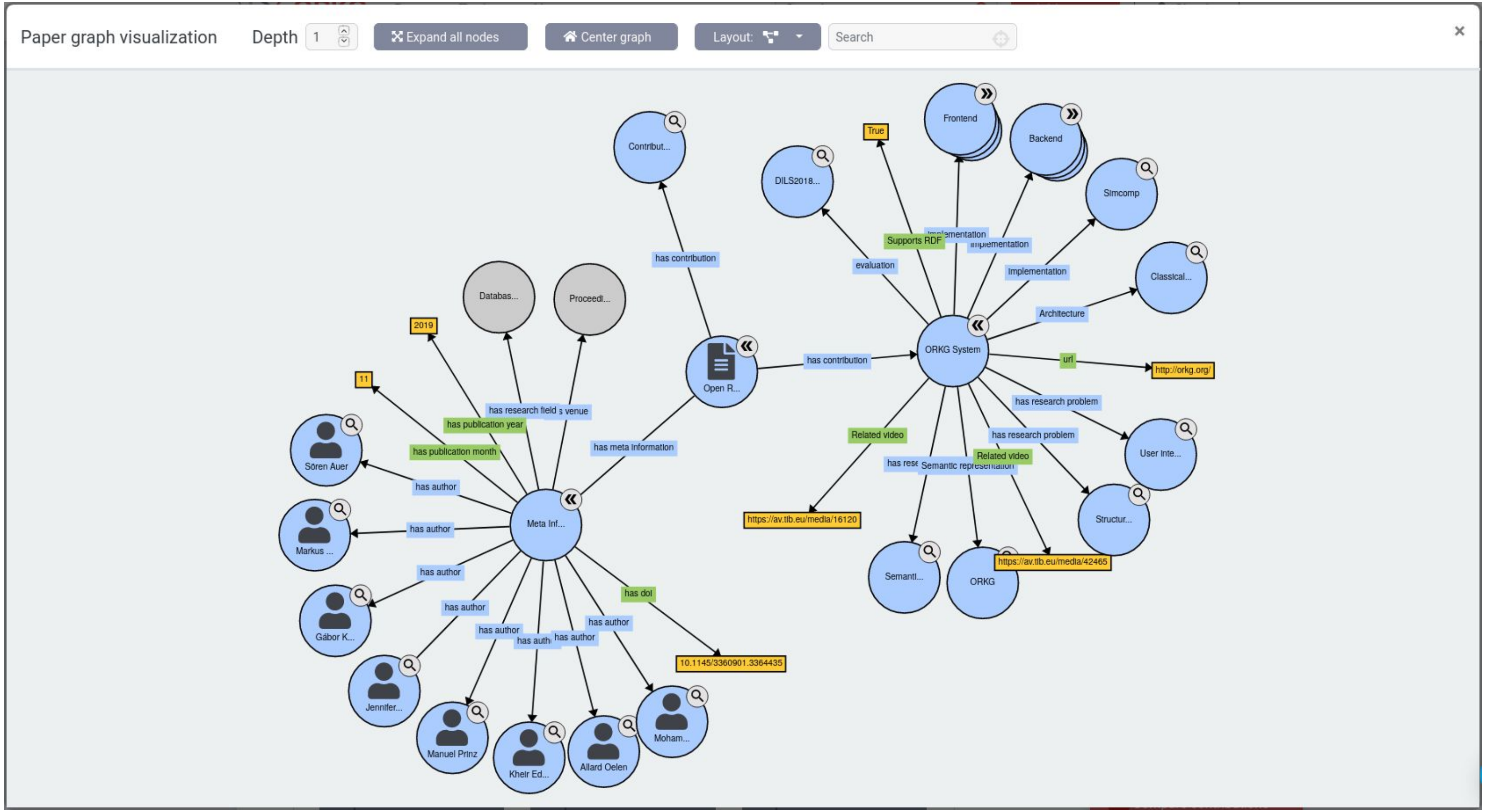
Is this it?

Currently **Goal**

And we still have a lot more (Abstract Annotator)



And we still have a lot more (Graph Visualizer)



And we still have a lot more (Tabular Data)

View dataset: DILS2018 User evaluation of ORKG frontend ×

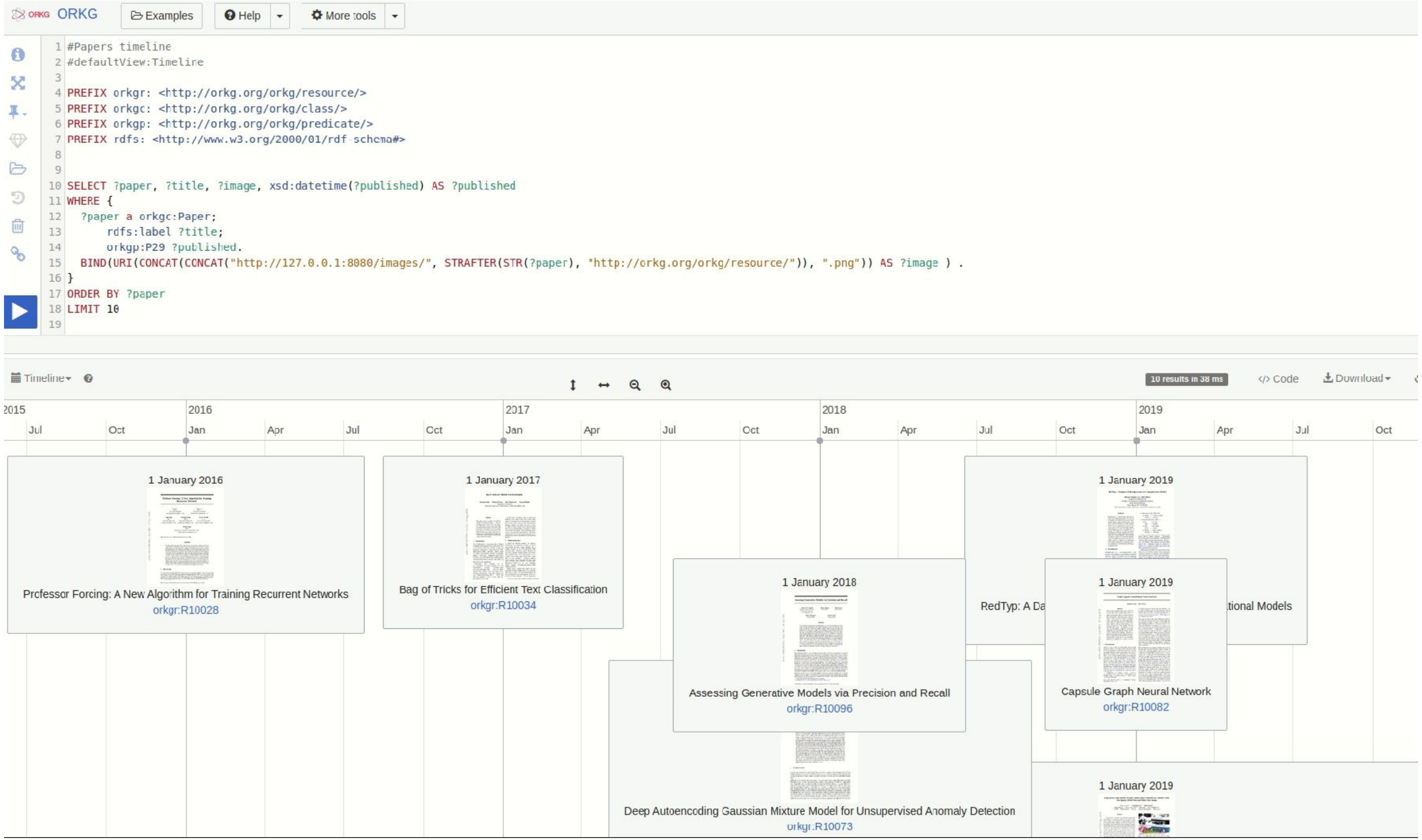
Showing 13 observations :

Options

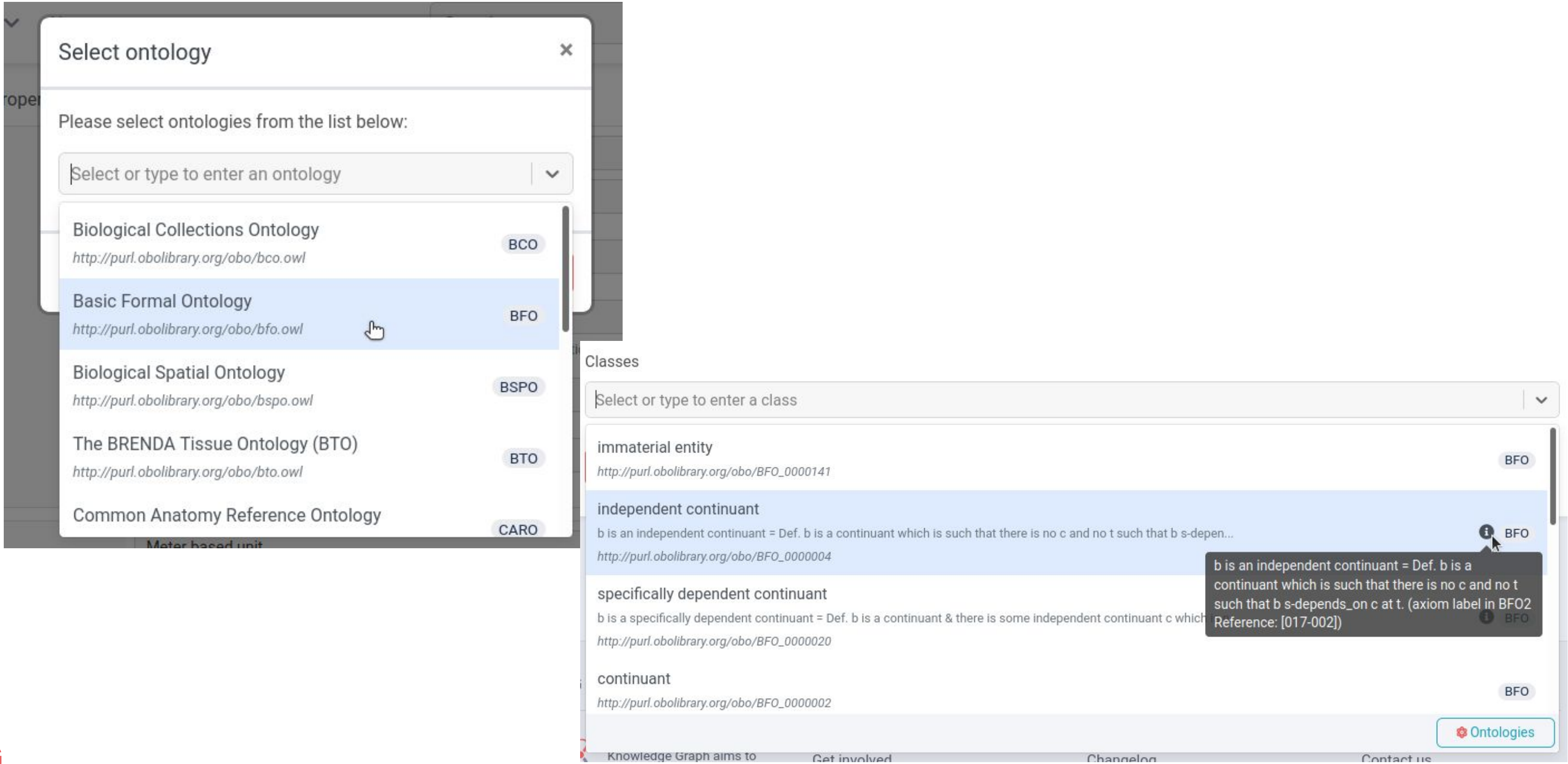
Participant Nr	Navigation (5=very intuiti...	Terminology (5=Easy to u...	Auto complete (5=very he...	Guidance needed (5=All t...	Suggest to Others (9=ver...	UI likeness (9=Very much)	Time (in mins)
Search...	Search...	Search...	Search...	Search...	Search...	Search...	Search...
1	4	4	5	3	2	6	16
2	2	3	5	4	8	7	19
3	4	5	5	3	9	7	15
4	3	3	5	3	6	7	13
5	4	3	5	3	6	8	14
6	4	3	5	3	8	9	13
7	3	4	5	3	7	6	19
8	3	2	4	3	8	6	13
9	4	5	3	3	7	5	14
10	4	5	5	1	8	8	22

Previous   Page 1 of 2   10 rows   Next

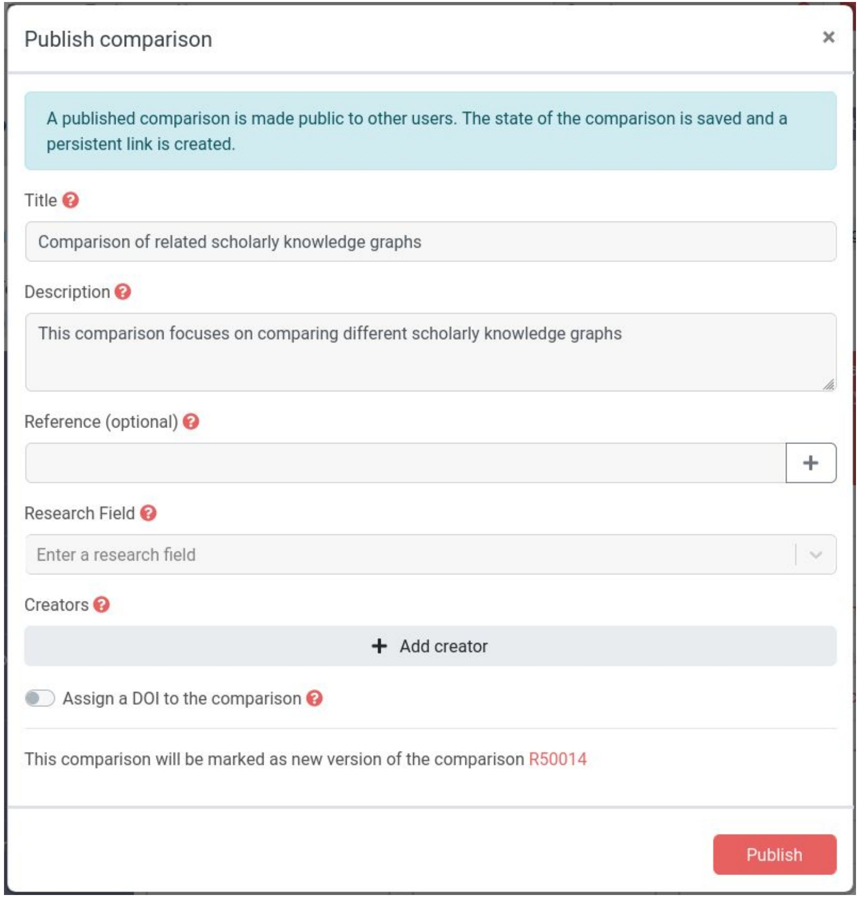
And we still have a lot more (RDF/SPARQL endpoint)



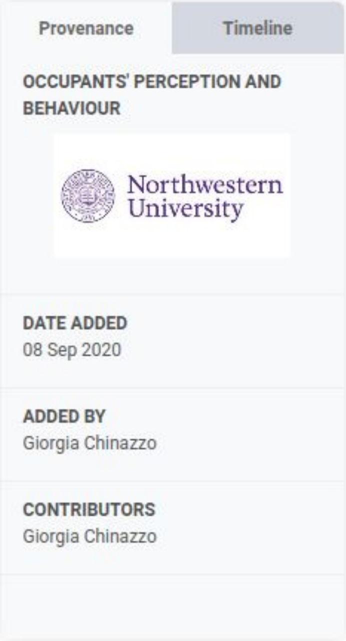
And we still have a lot more (Linking to external ontologies)

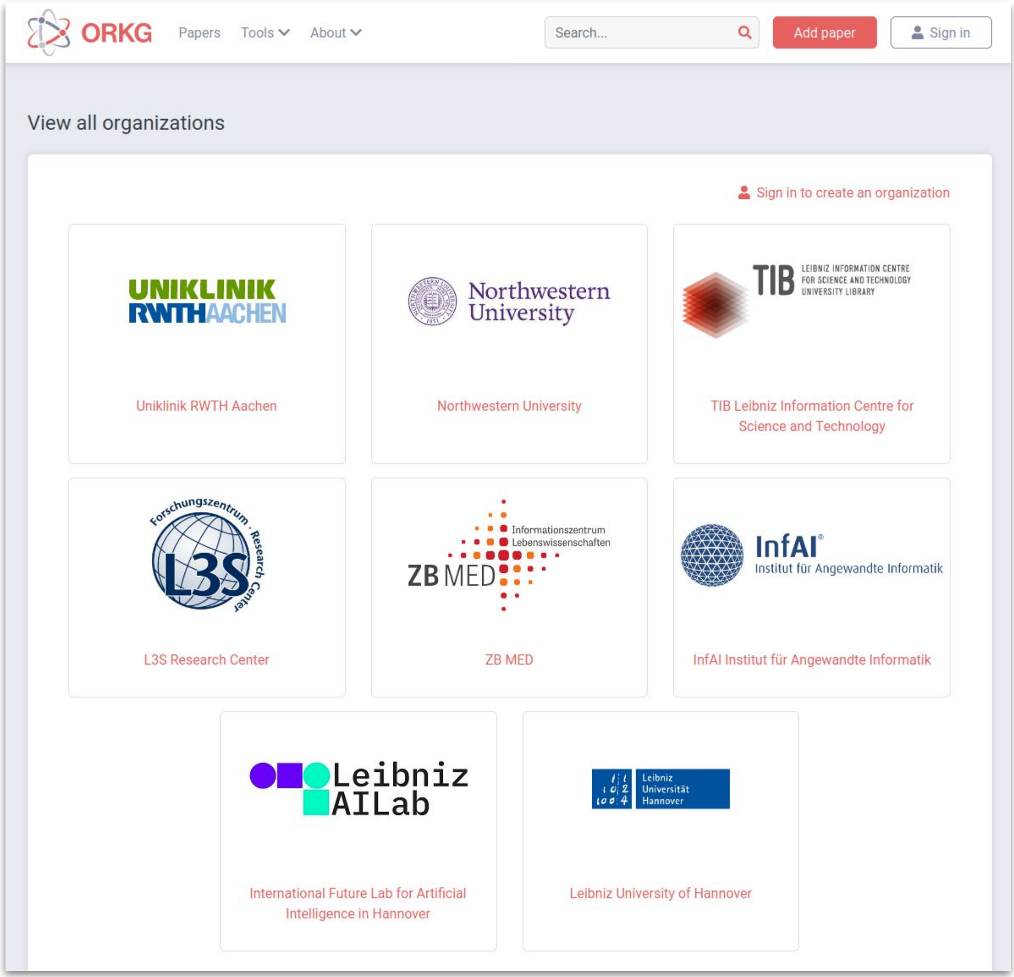


And we still have a lot more (Publishing Comparisons)

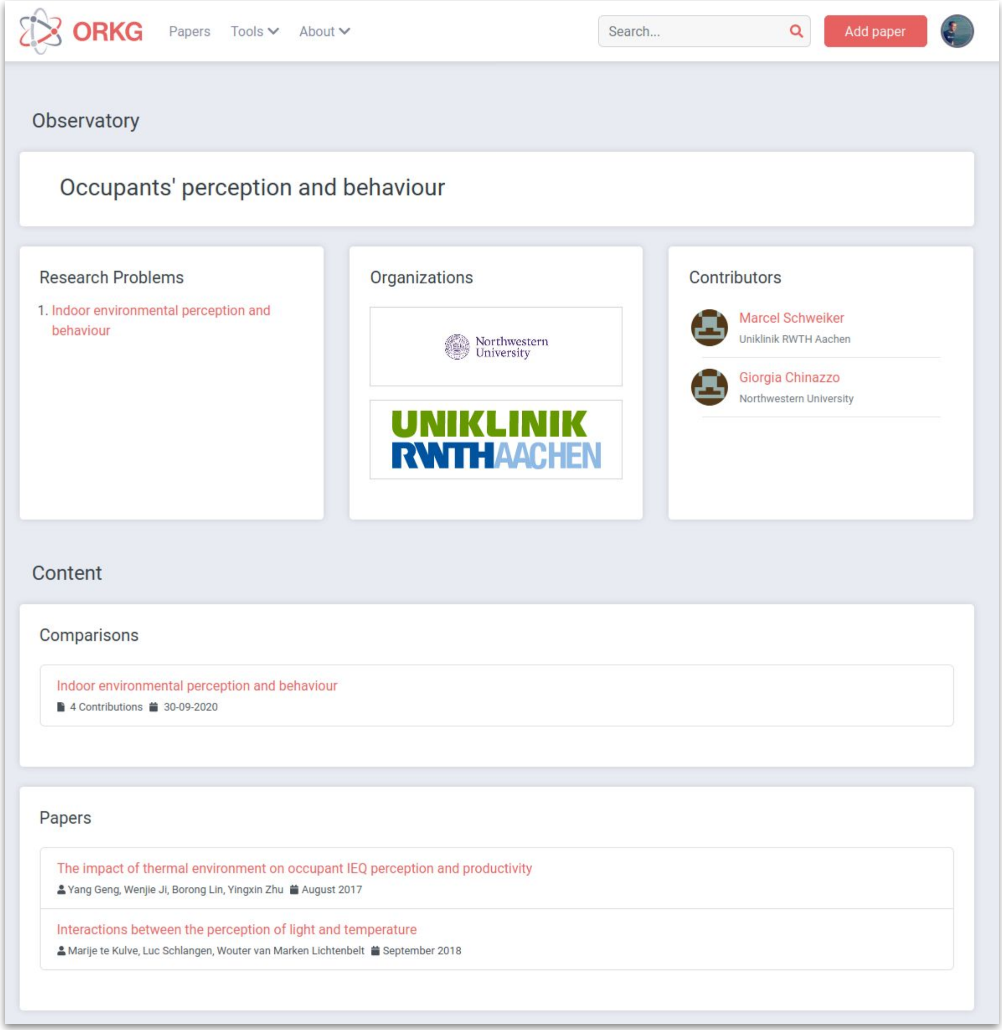


And we still have a lot more (Organizations)





And we still have a lot more (Observatories)



behind the scenes

The Awesome Team behind the ORKG

Sören Auer
Lead

Vitalis Wiens
PhD Student

Muhammad Haris
PhD Student

Jennifer D'Souza
Researcher

Mohamad Yaser Jaradeh
Developer

Golsa Heidari
PhD Student

Lars Vogt
Researcher

Markus Stocker
Co-Lead

Arthur Brack
PhD Student

Allard Oelen
Developer

Kheir Eddine Farfar
Developer

Manuel Prinz
Developer

And many more ….

Any questions?

Get your hands dirty now: orkg.org

Yaser Jaradeh (jaradeh@l3s.de)

Open Research Knowledge Graph