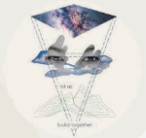
A DECADE OF SCHOLARLY RESEARCH ON OPEN KNOWLEDGE GRAPHS

Houcemeddine Turki, Abraham Toluwase Owodunni, Mohamed Ali Hadj Taieb, René Fabrice Bile, Mohamed Ben Aouicha

LREC-COLING 2024 – Turin, Italy, 20-25 May 2024

Acknowledgments



MASAKHANE

Support

VILÉM ZOUHAR

Data Visualization and Analysis

WIKIMEDIA FOUNDATION

Funding

SISONKEBIOTIK

Support

Data Engineering and Semantics

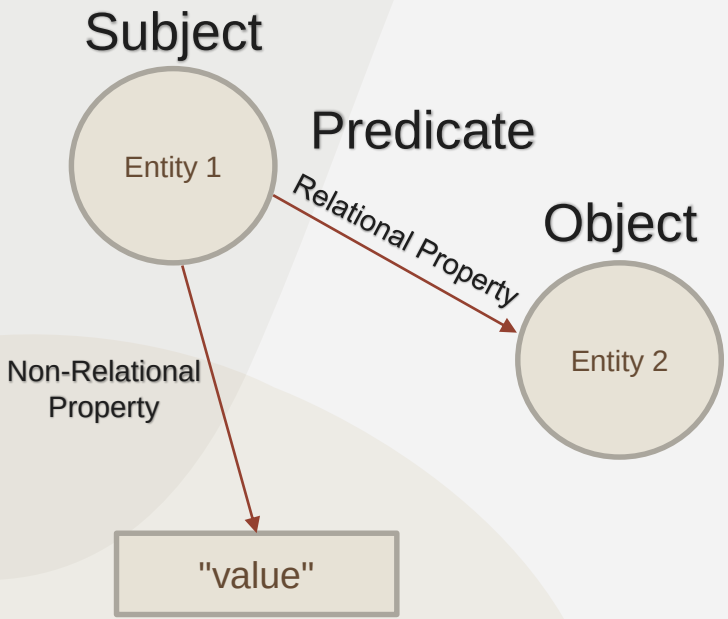
- A research unit at the University of Sfax, a top-tier Tunisian public university at the top of computer science rankings in Africa.
- Specialized in Data and Knowledge Engineering, Social Network Analysis, and Open Science.
- Has a data center and is headquartered within the Faculty of Sciences of Sfax.

https://deslab.org

mohamed.benaouicha@fss.usf.tn

@DES_Unit

Open Knowledge Graphs



- **F**indable, **A**ccessible, **I**nteroperable, and **R**eusable.
- Organizes the information in the form of triples (Subject – Predicate – Object).
- Readable by human and machine.

WIKIDATA

DBpedia

yago
select knowledge

BabelNet

INTRODUCTION

01

INCREASED INTEREST

Open Knowledge Graphs are advocated as valuable databases

02

RESOURCES CREATED

Many Lexico-Semantic Open Knowledge Graphs have been created

03

AWARENESS NEEDED

Natural Language Processing Community needs to know this

04

UNASSESSED FIELD

The effect of Open Knowledge Graphs on research is not quantified

METHODS

01

DATA COLLECTION

Scopus®

We retrieve scholarly publications from Scopus

02

DATA FILTERING

We isolate papers explicitly dedicated to open knowledge graphs

03

STATISTICAL ANALYSIS

matplotlib

We quantify the publication count and citation impact of publications according to various criteria

04

NETWORK ANALYSIS

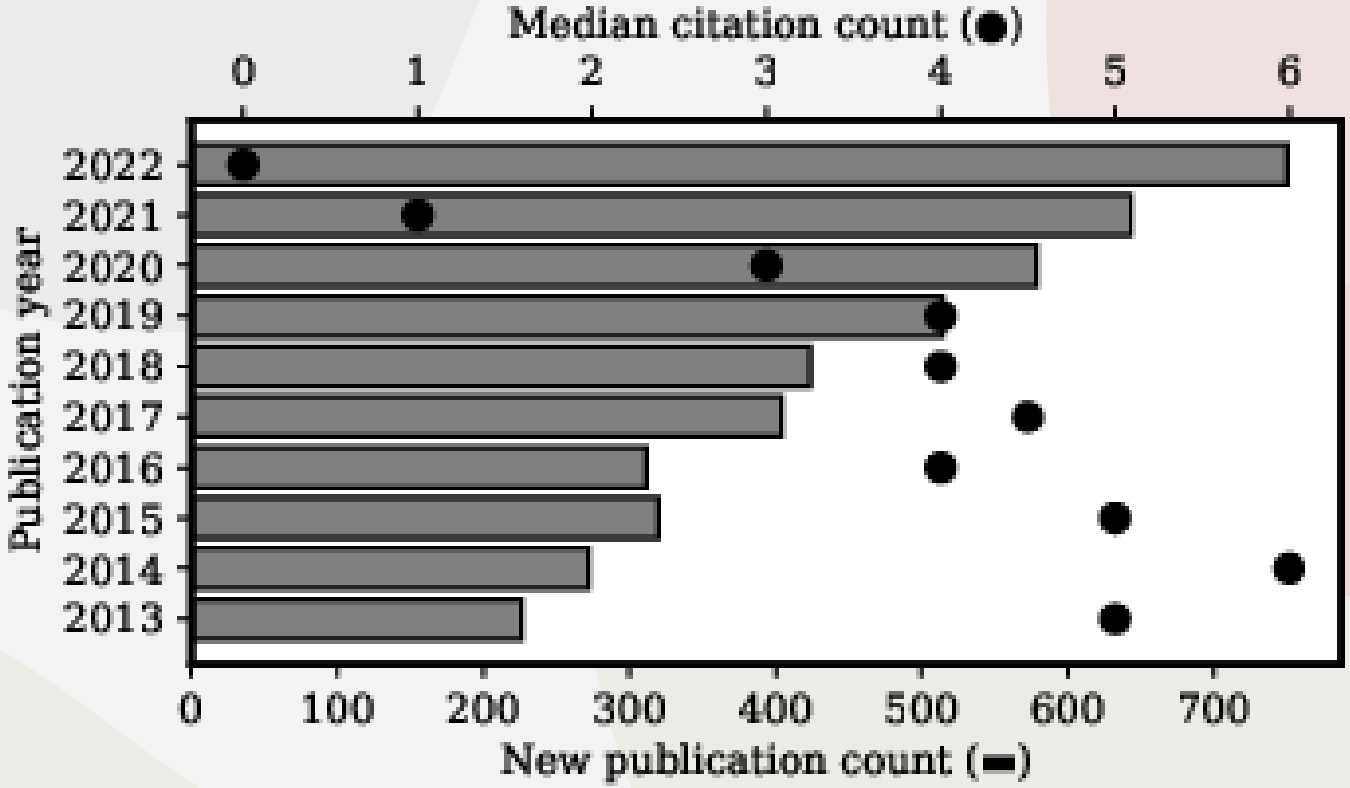
VOSviewer
Visualizing scientific landscapes

We generate the keyword co-occurrence networks for five periods between 2013 and 2022

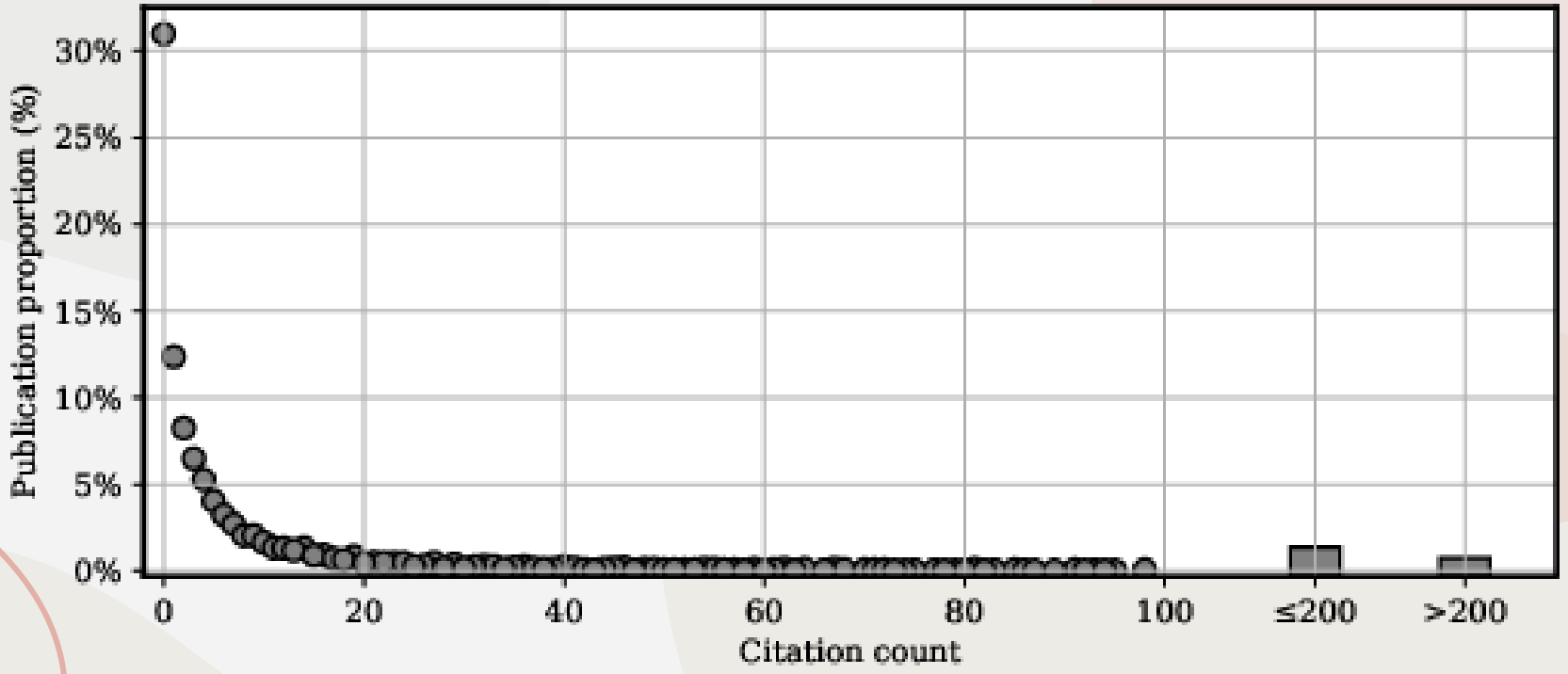
RESULTS AND DISCUSSION

What you should know about open knowledge graph research

Publications per year



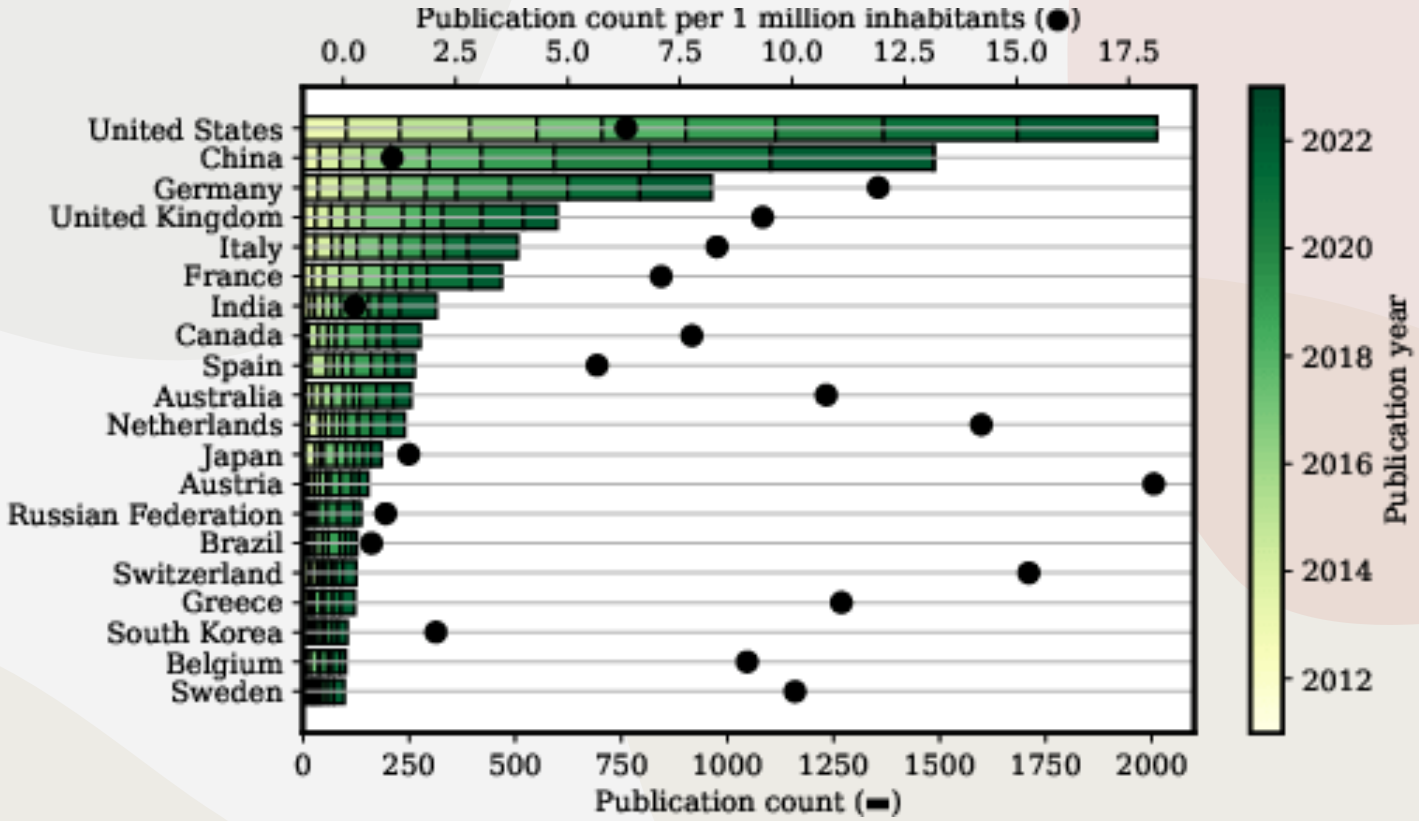
Citation Distribution



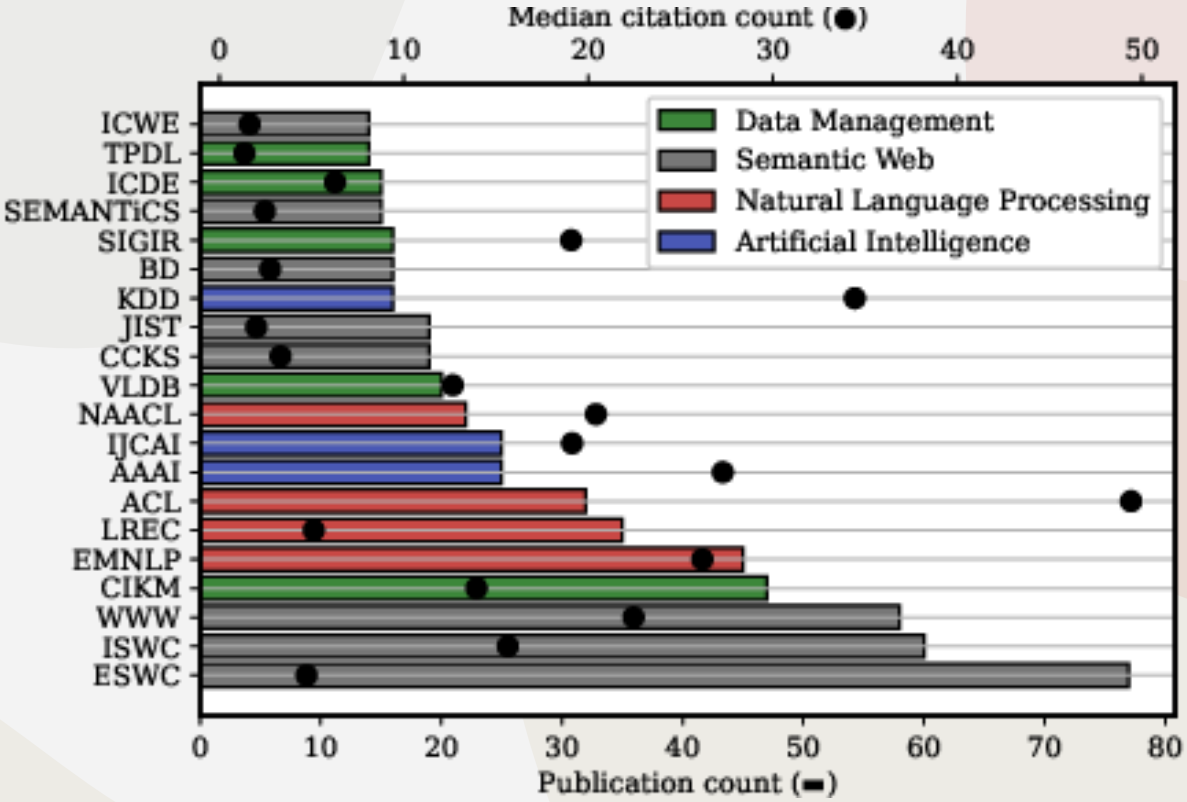
Most Cited Research Papers

Rank	Paper	Cited
1	**Sharing and community curation of mass spectrometry data with Global Natural Products Social Molecular Networking** Mingxun Wang, et al. (2016) Nature Biotechnology, 34(8):828– 837.	1854
2	**DBpedia – A large-scale, multilingual knowledge base extracted from Wikipedia** Jens Lehmann, et al. (2015) Semantic Web, 6(2):167–195.	1715
3	**Wikidata: a free collaborative knowledgebase** Denny Vrandečić and Markus Krötzsch (2014) Communications of the ACM, 57(10):78–85.	1617
4	**ChestX-ray8: Hospital-Scale Chest X-Ray Database and Benchmarks on Weakly Supervised Classification and Localization of Common Thorax Diseases** Xiaosong Wang, et al. (2017) Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pages 3462–3471. IEEE.	1541
5	**The Reactome pathway knowledgebase** David Croft, et al. (2014) Nucleic Acids Research, 42(D1):D472–D477.	1176
6	**The IUPHAR/BPS Guide to PHARMACOLOGY: an expert-driven knowledgebase of drug targets and their ligands** Adam J. Pawson, et al. (2014) Nucleic Acids Research, 42(D1):D1098–D1106.	777

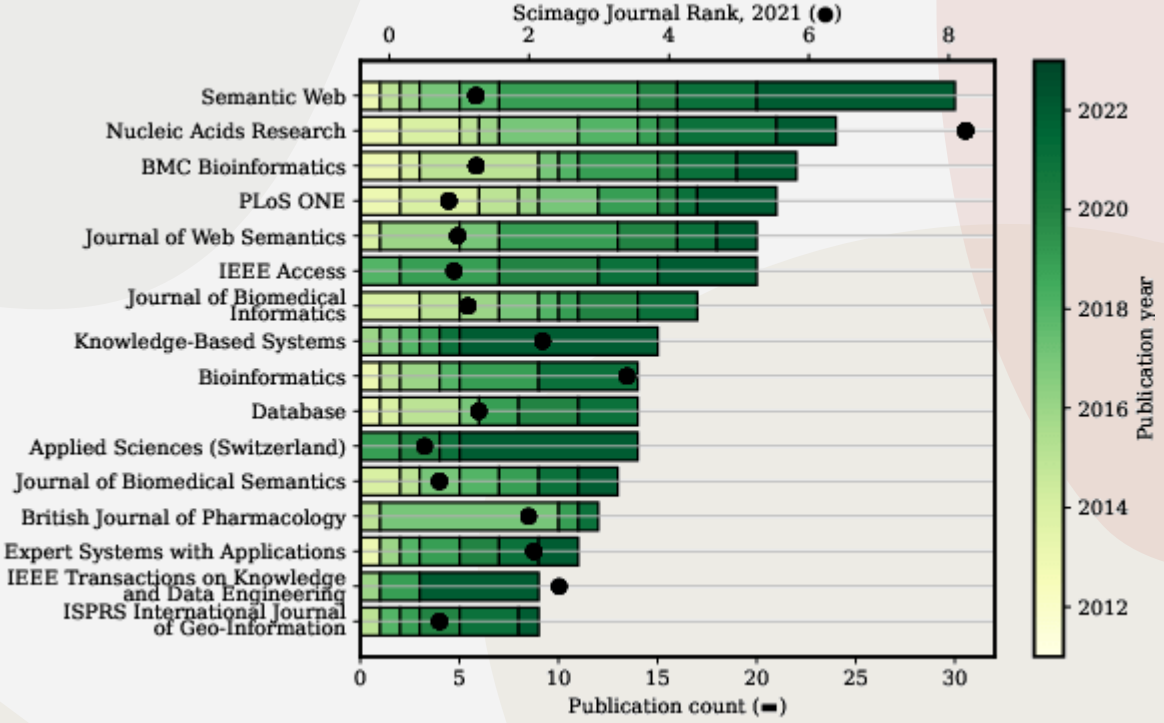
Most Published Countries



Most Published Conferences

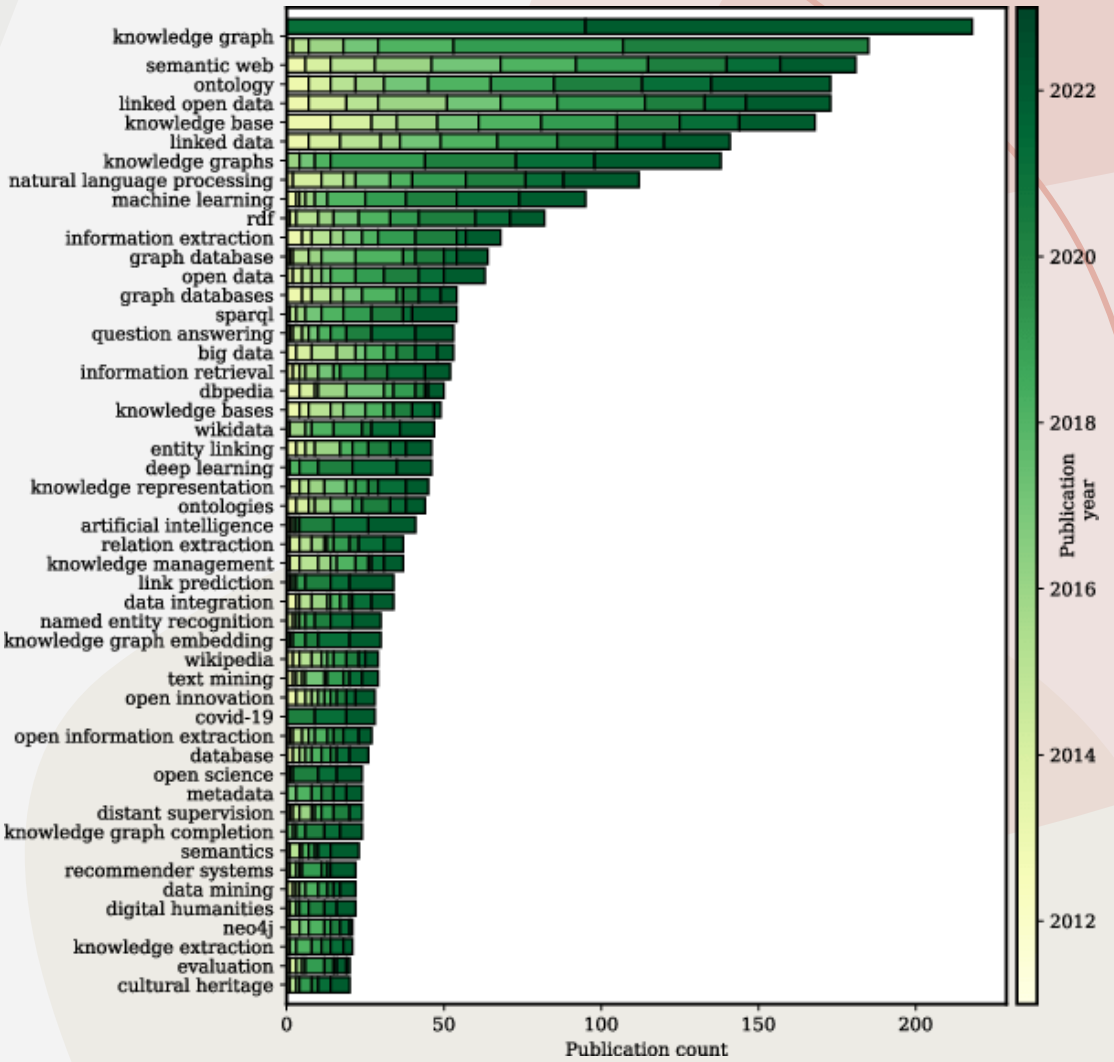


Most Published Journals

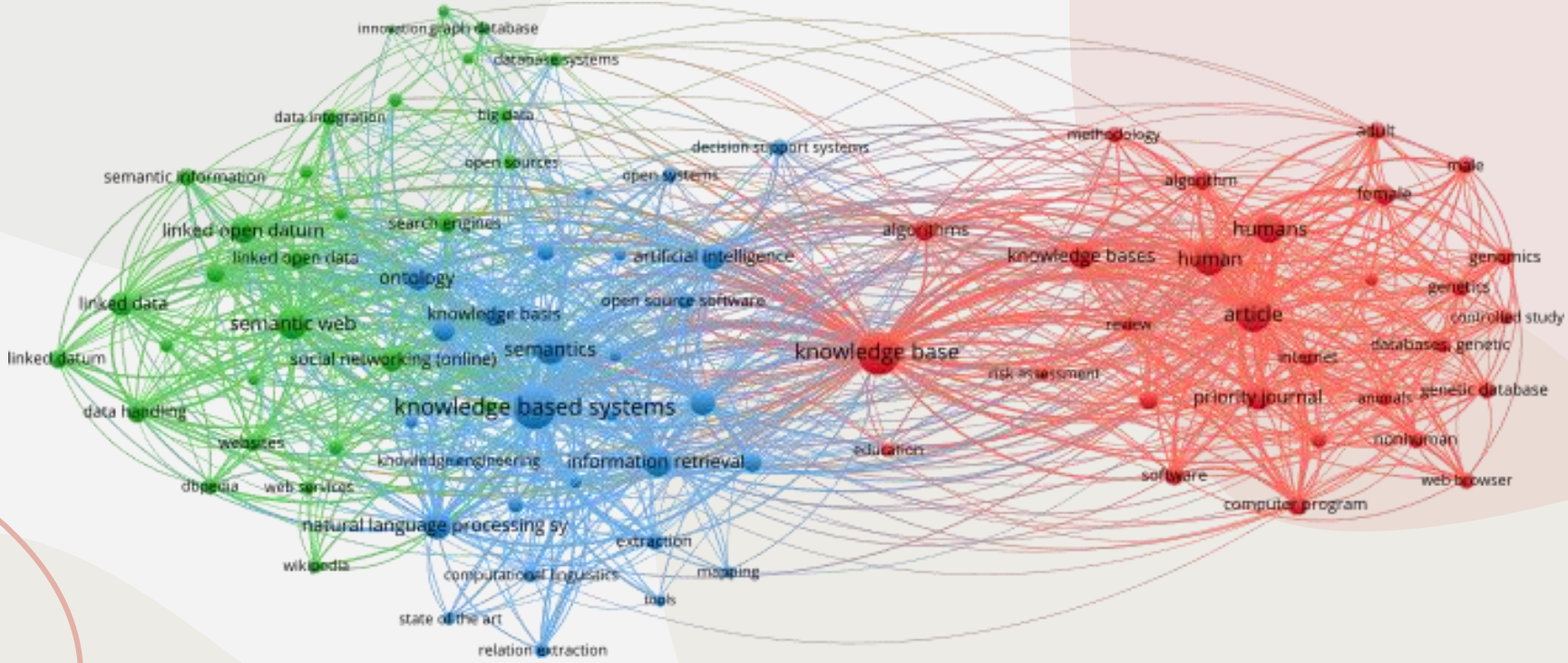


Top Keywords

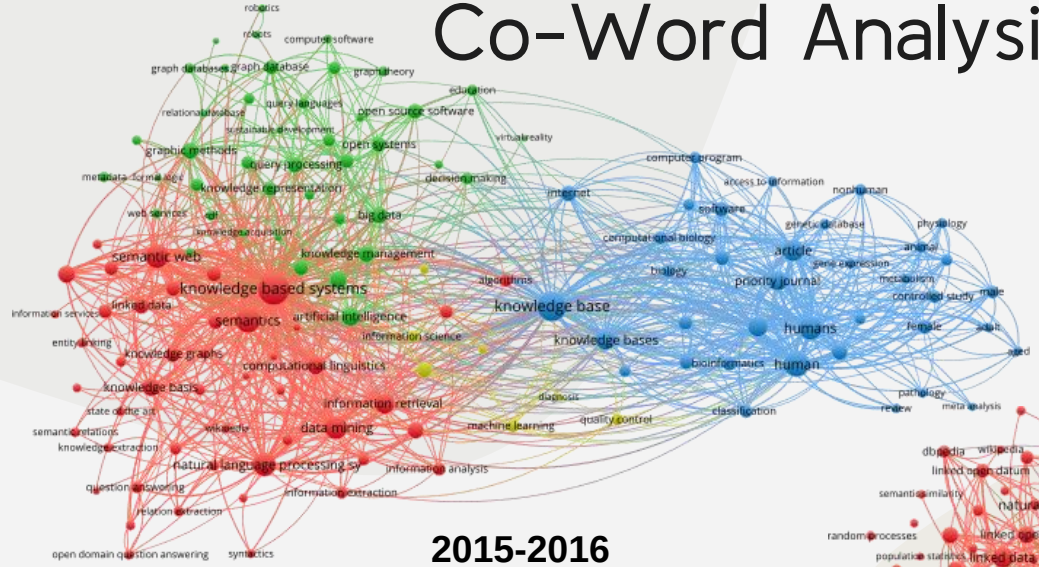
- The **basic terminology** of the **semantic web** field is featured during all the period between 2013 and 2022.
- The **first period** (2013-2016) is characterized by works on the **creation** and **querying** of open knowledge graphs.
- The **second period** (2017-2019) defines **multidisciplinary applications** of open knowledge graphs using **deep learning** and **embeddings**.
- The third period (2020-2022) urges the move towards **open science** and the **mass development** and **scalability** of open knowledge graphs.



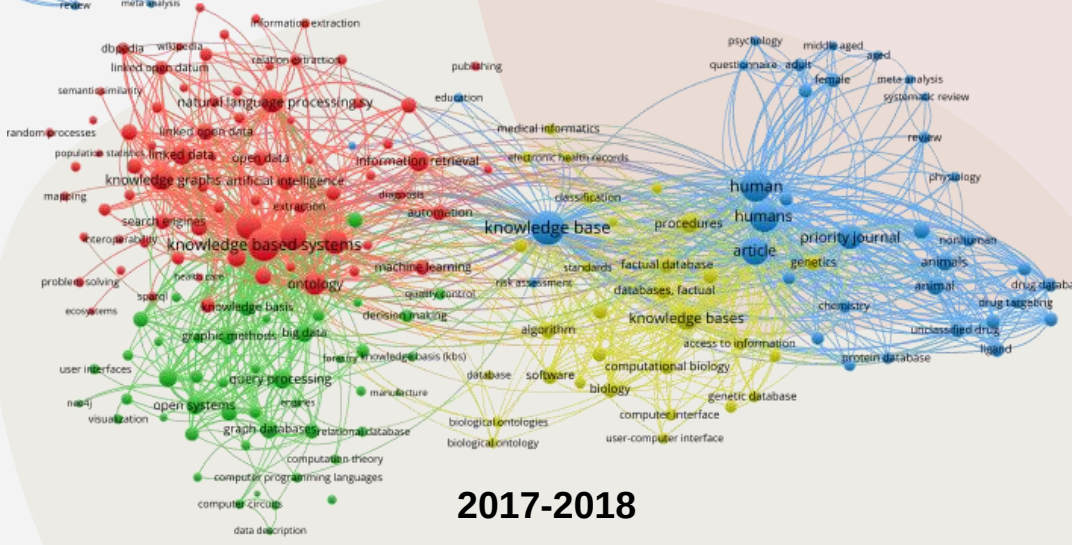
Co-Word Analysis: 2013-2014



Co-Word Analysis: 2015-2018

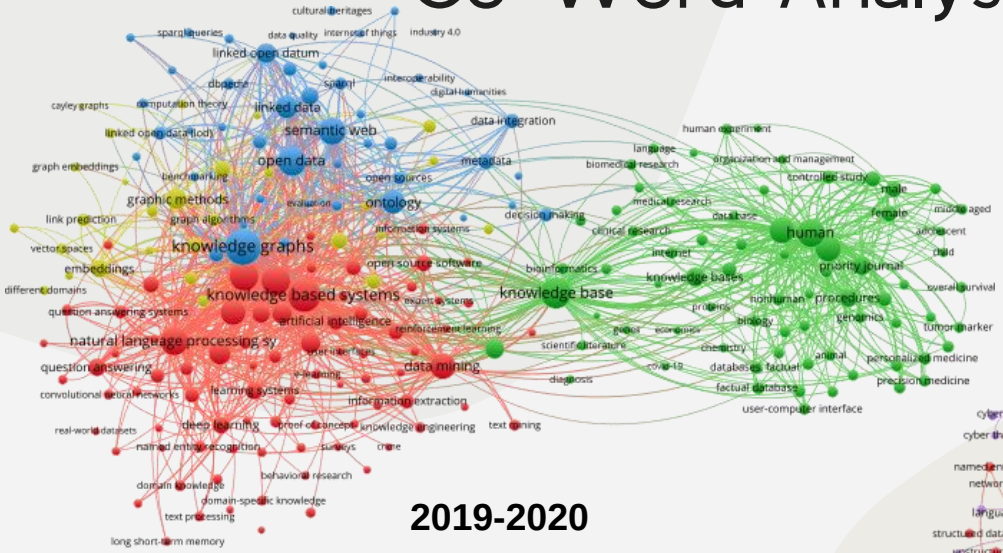


2015-2016

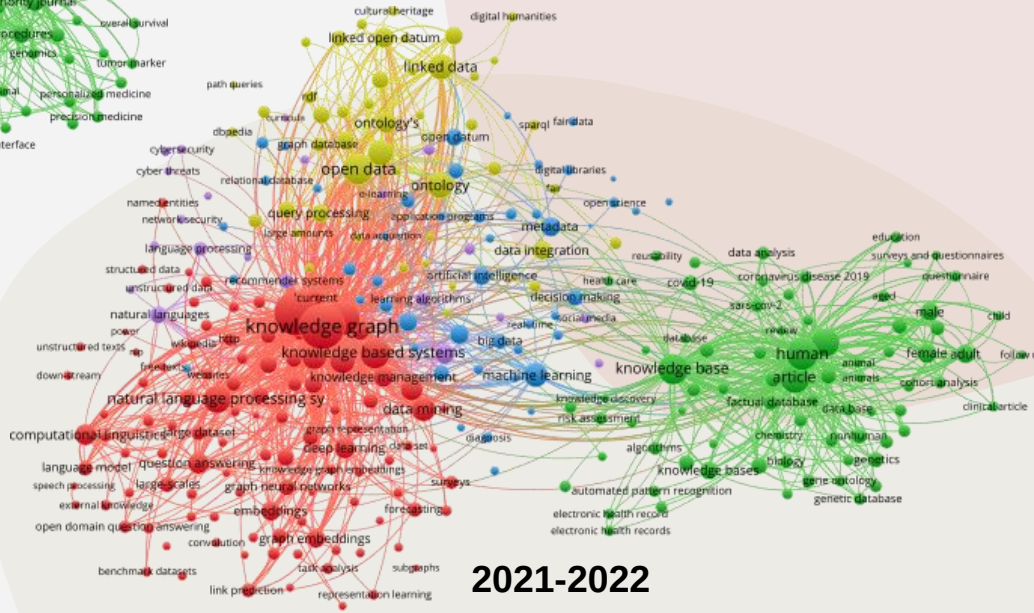


2017-2018

Co-Word Analysis: 2019-2022



2019-2020



2021-2022

CONCLUSION

- Research on **open knowledge graphs** expanded a lot as **new concepts** keep emerging.
- **Development** is ahead of **Research**.
- A **triple helix relation** between government, research, civil society, and industry is required.
- Encouraging **cross-disciplinary research** is needed.
- A **more detailed analysis** of the topic is useful.

Thank you

Houcemeddine Turki.
turkiabdelwaheb@hotmail.fr
@Csisc1994

AR FR EN IT