



#### THE NETWORKS OF SCIENCE. DATA-DRIVEN UNDERSTANDING OF SCIENTIFIC PRODUCTION. Diego Kozlowski FSTM

5/6 March,2020

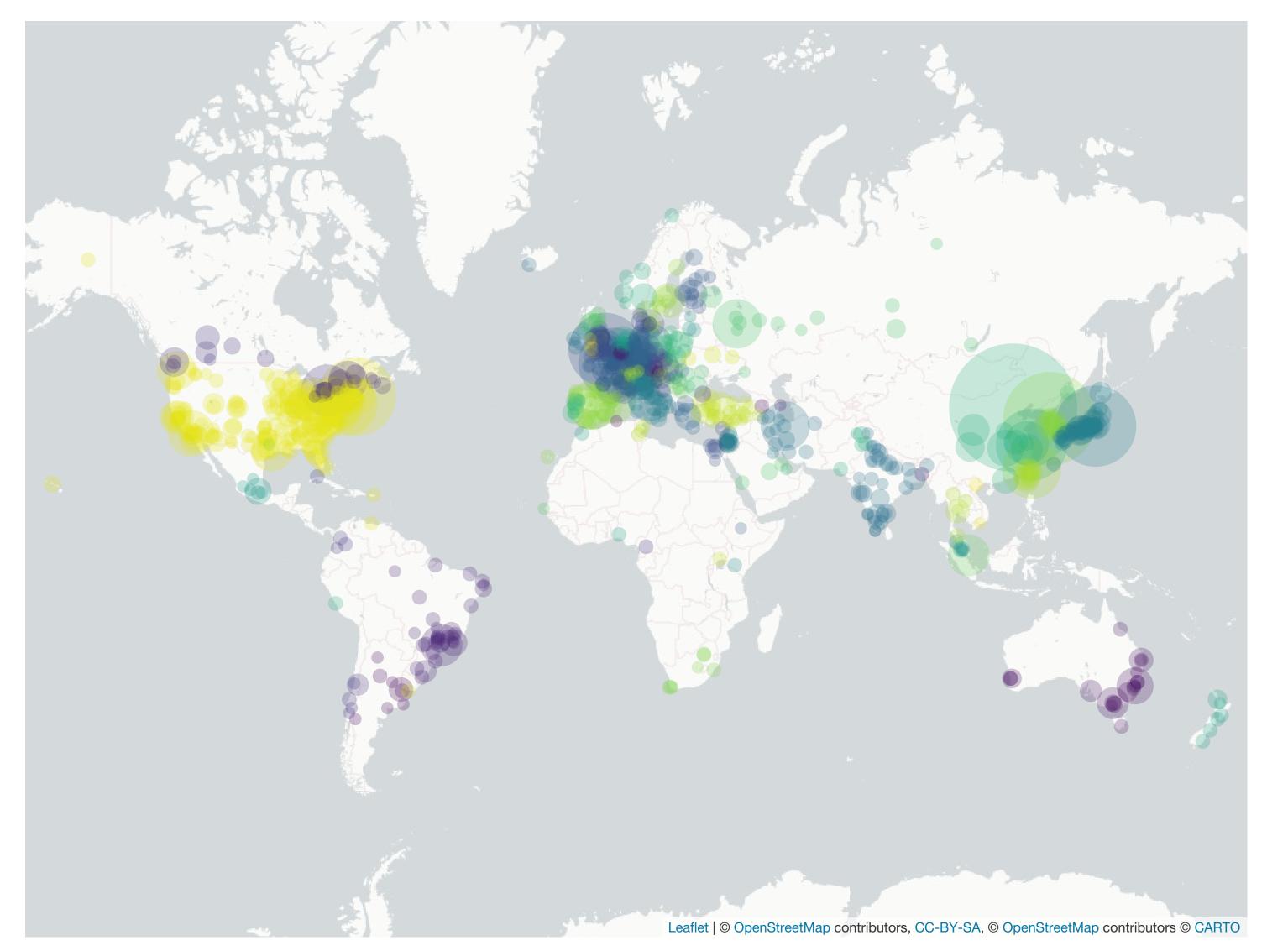
# AIM AND SCOPE

- Science of Sciences is a developing field that focus on quantitative studies of scientific production.
- As a social activity, scientific output is not independent from the society in which it is produced.
- This project aims to understand how science is shaped by society and vice-versa.

## OBJECTIVES

Advance in the understanding of Science's development across *fields* and *countries*,

#### FIGURE 3



in our current society, from a *data-driven*, *epistemic* and *social* perspective.

- **1** Analyze the spread of new ideas within a field using semantic analysis,
- 2 Study the spread of knowledge across fields, re-conceptualize the notion of *basic* science based on the evidence found in data of supplier and demanding fields,
- 3 Analyze the impact of collaboration between authors, institutions, cities, countries. Understand if there are underlying inequalities on these different hierarchical levels, where the rich get richer [1],
- Image: Image: A constraint of the second study career paths across is a constraint countries and fields. Explore possible inequalities related to gender and ethnicity,
- **5** Extend the concepts of *Globalized Science* and *Knowledge Economy* to the different roles countries play in the international production of science, and its relation with the role these play in global economy as a whole;

## METHODOLOGY

There are two main methodologies to be use:

- Graph theory, to study the links between science practitioners (researchers, institutions, countries),
- Natural Language Processing, to do large-scale analysis over scientific communications (papers).

# FIGURE 1

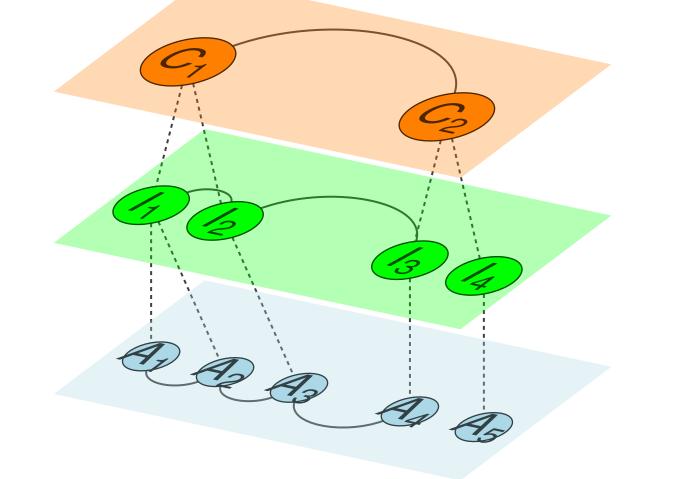
## FIGURE 2

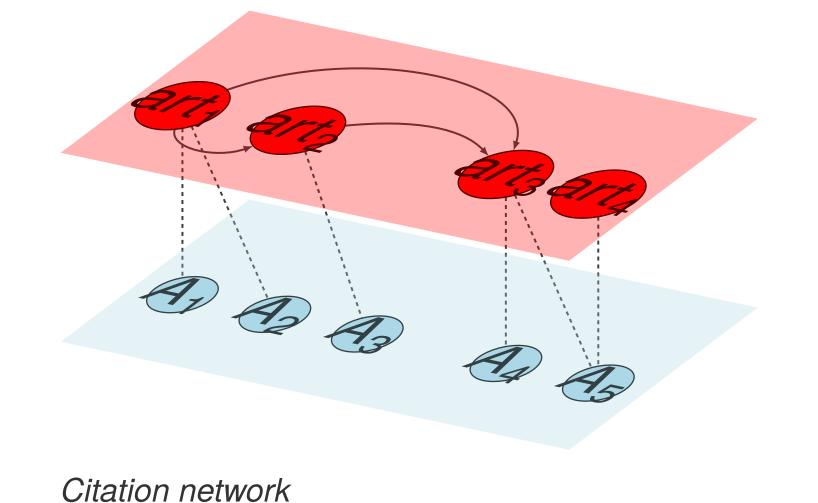
Top 1000 scientific producers on city level. 2010



for evolution of cities, visit: FIGURE 4

#### Proportion of women by field Germany

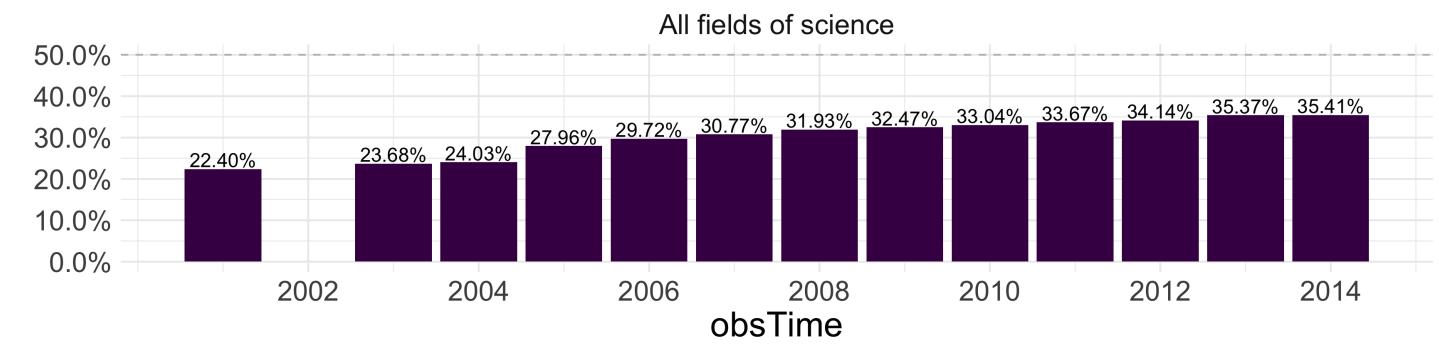


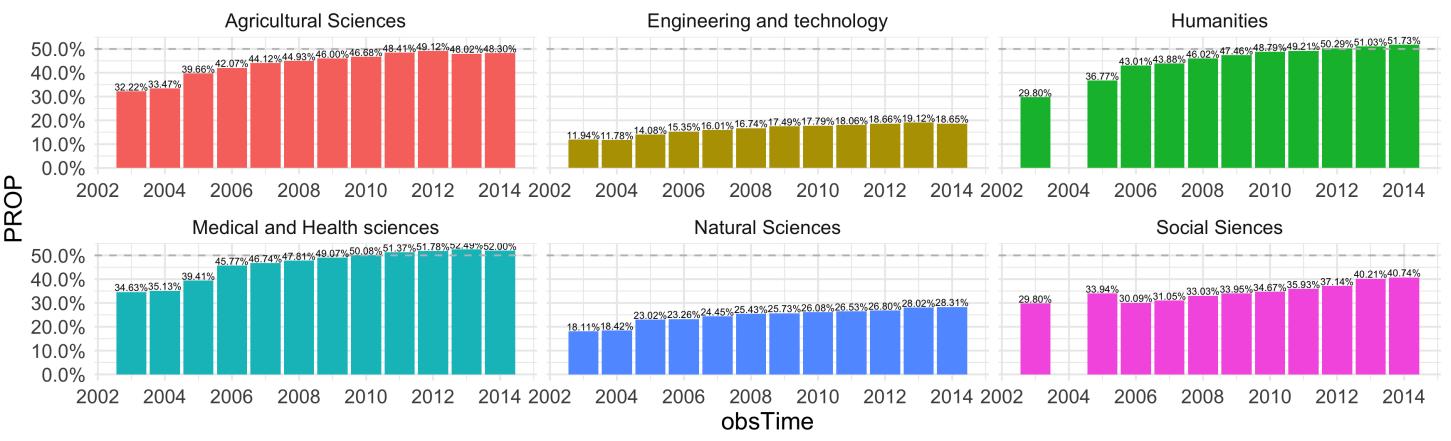


Collaboration network

# WORK PACKAGES

- **1** Literature review
  - Science of Science, Sociology of research, Social Network Analysis and Natural Language Processing,
  - Develop an automatic literature review based on NLP;
- 2 Data collection and assimilation
  - Implement a name disambiguation over Word of Science database
  - collect and process alternative sources of data;
- **3** Modeling
  - Build a hierarchical multi-level network on author-institution-city-country. Study the topological patterns of these networks, both defining links as collaborations, and as a bipartite graph of articles and authors,





Female researchers across fields. Germany. 2010

## ACKNOWLEDGEMENT

The Doctoral Training Unit **Data-driven computational modelling and applications** (DRIVEN) is funded by the Luxembourg National Research Fund under the PRIDE programme (PRIDE17/12252781). https://driven.uni.lu

- study the citation map across disciplines,
- analyze the career path differences across countries and fields. Compare the position of institution in the collaboration network with the researchers movement patterns,
- aggregate the conclusions on country level and analyze the differences in scientific production across countries.

#### REFERENCES

[1] Roger King. "Power and networks in worldwide knowledge coordination: The case of global science". In: *Higher* Education Policy 24.3 (2011), pp. 359–376. ISSN: 09528733. DOI: 10.1057/hep.2011.9. URL: http://dx.doi.org/10.1057/hep.2011.9.