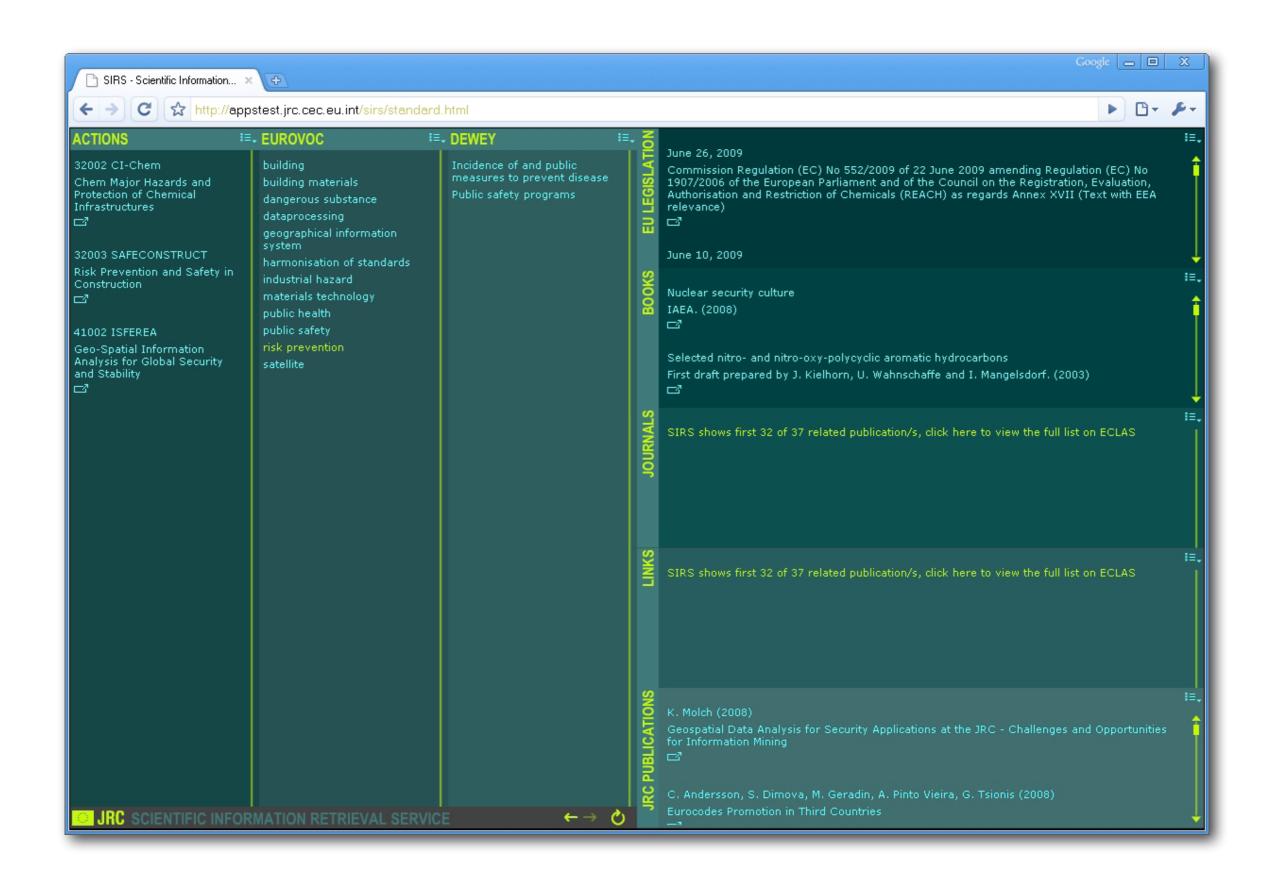
# Using Networks to Visualize Publications

#### Dario Rodighiero

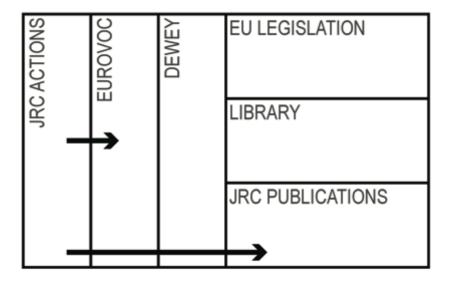
Digital Humanities Laboratory (DHLAB) School of Architecture, Civil and Environmental Engineering (ENAC), College of Humanities (CDH) École polytechnique fédérale de Lausanne (EPFL)

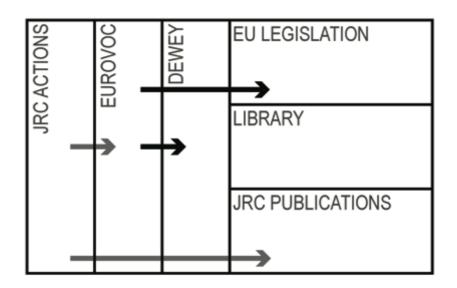


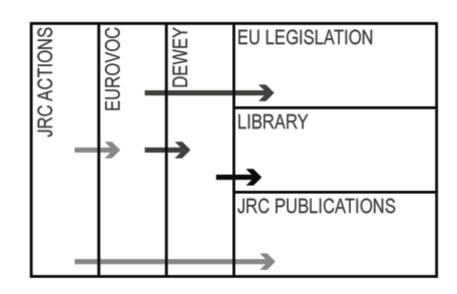
JRC Scientific Library and Carlo Ferigato in 2009

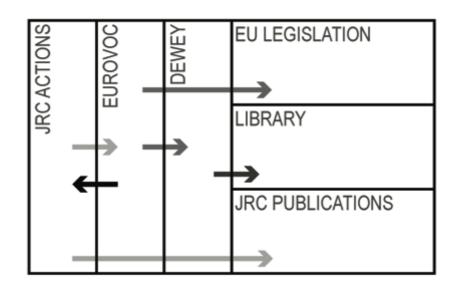


Role of Thesauri in a Scientific Organisation (Ferigato, Merlo, Panfili, and Rodighiero 2008)









Visualizing Using Lists

"the [...] list can be exemplified [...] by a library catalogue [...] and even by a dictionary that records all the words in the lexicon of a given language."

"the archive, as distinct form of collection or library, constitutes a repository or ordered system of documents and records, both verbal and visual [...]"



The Golden Stairs, Tate Gallery, London (Edward Burne-Jones 1880)

1 Principle of Later in Time

2 Principle of Later in Evolution

3 Principle of Spatial Contiguity

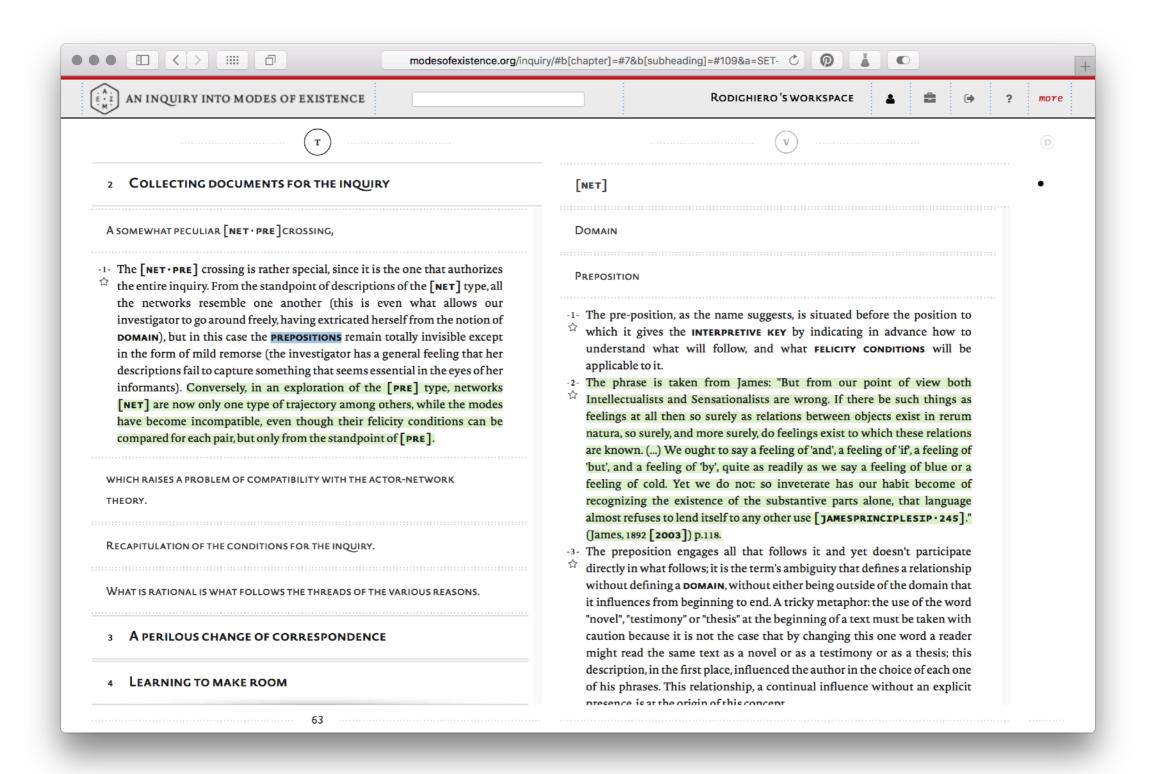
4 Principle of Quantitative measure

5 Principle of Increasing Complexity

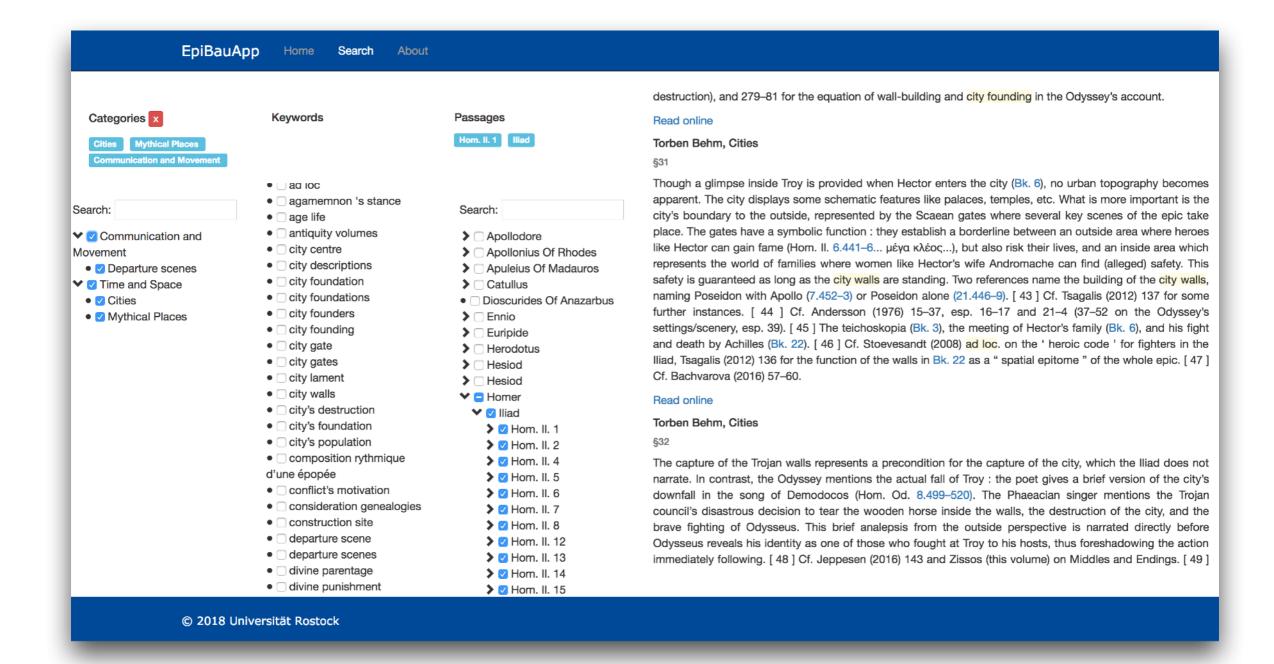
6 Principle of Canonical Sequence

7 Principle of Literary Warrant

8 Principle of Alphabetical Sequence

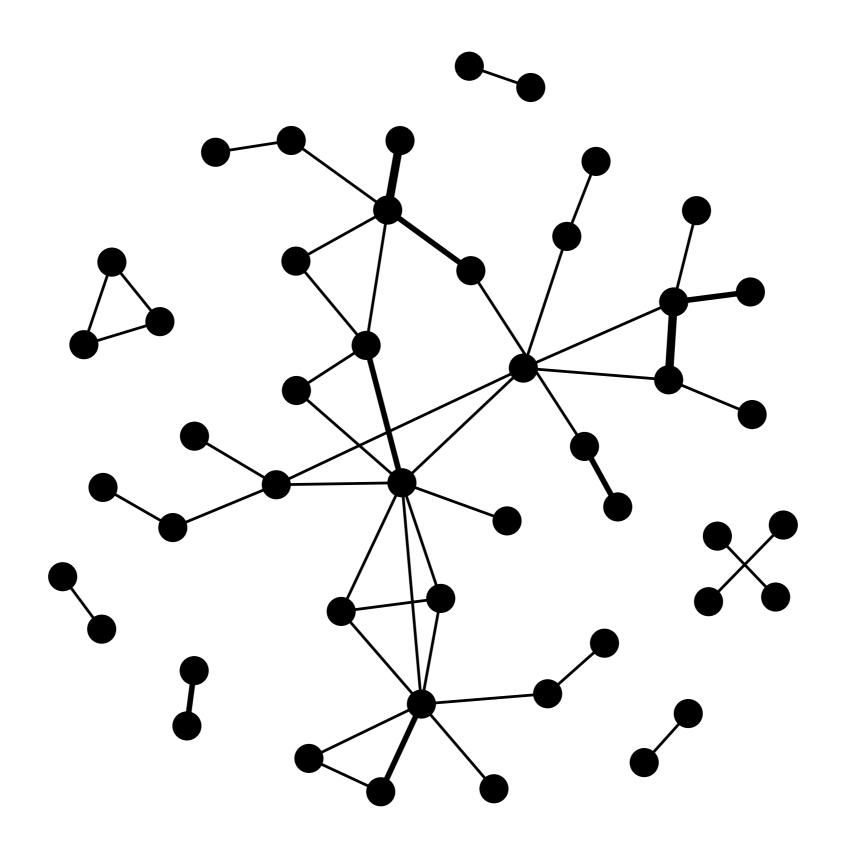


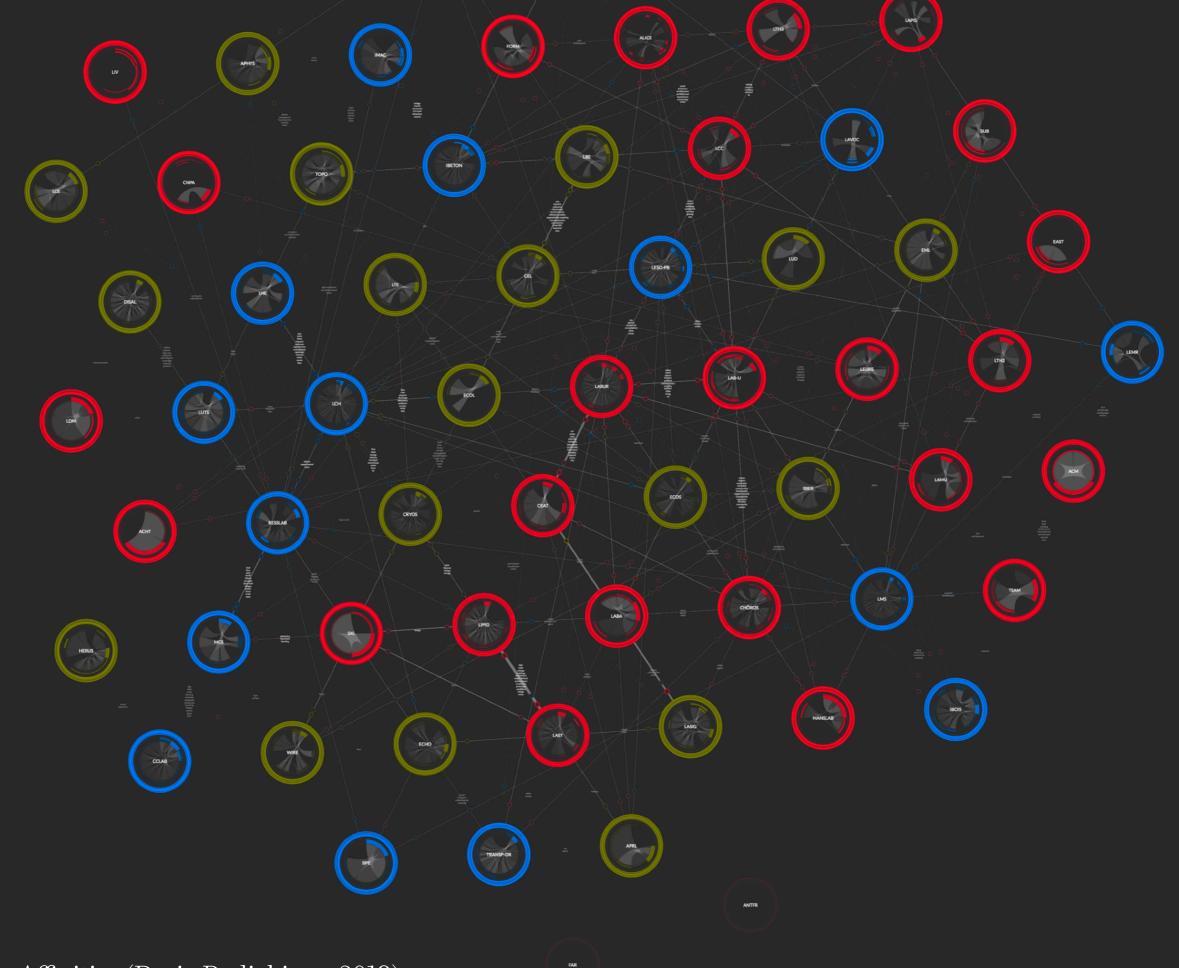


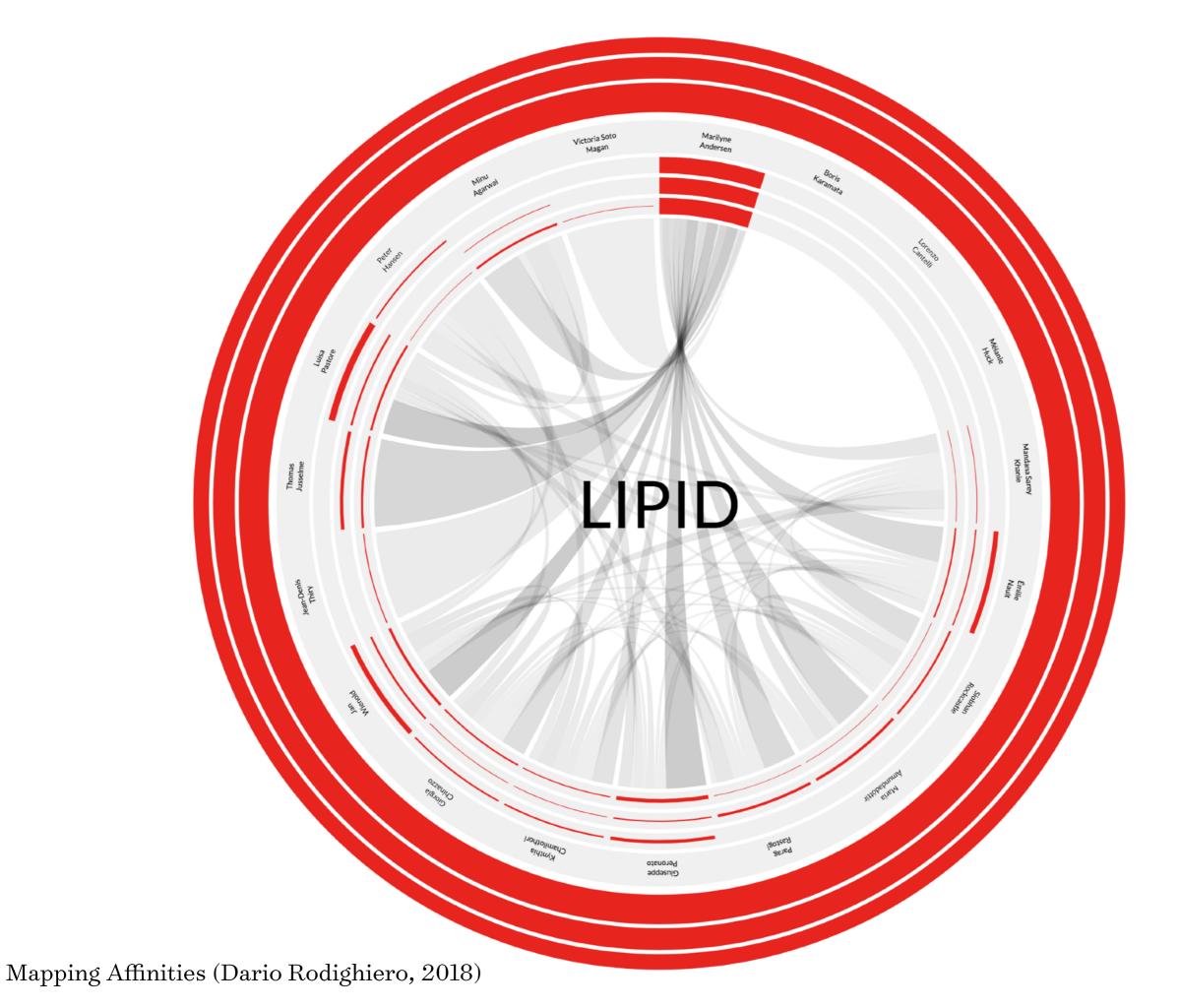


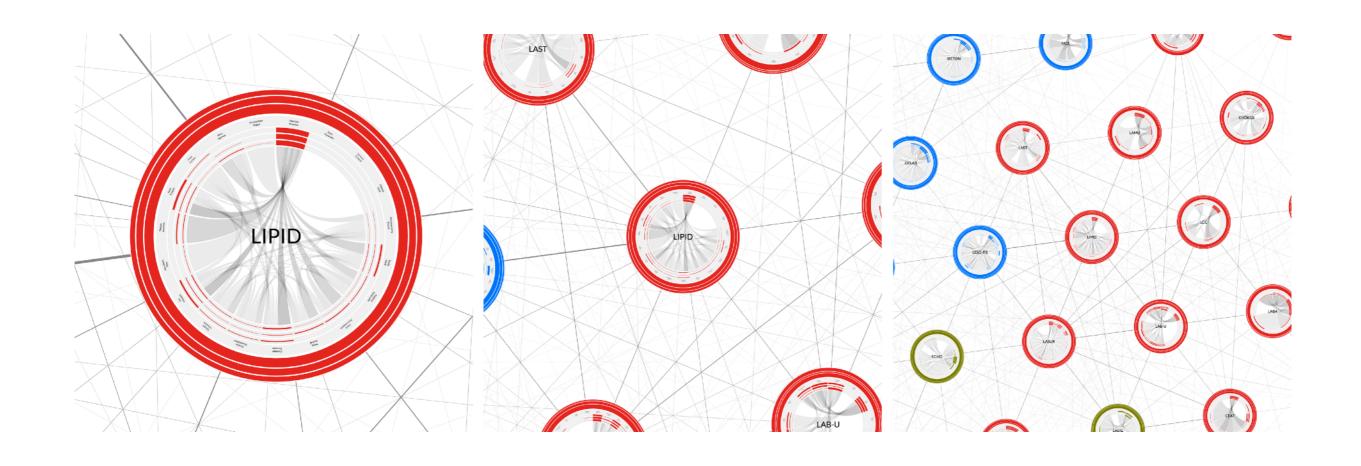
Visualizing Using Networks

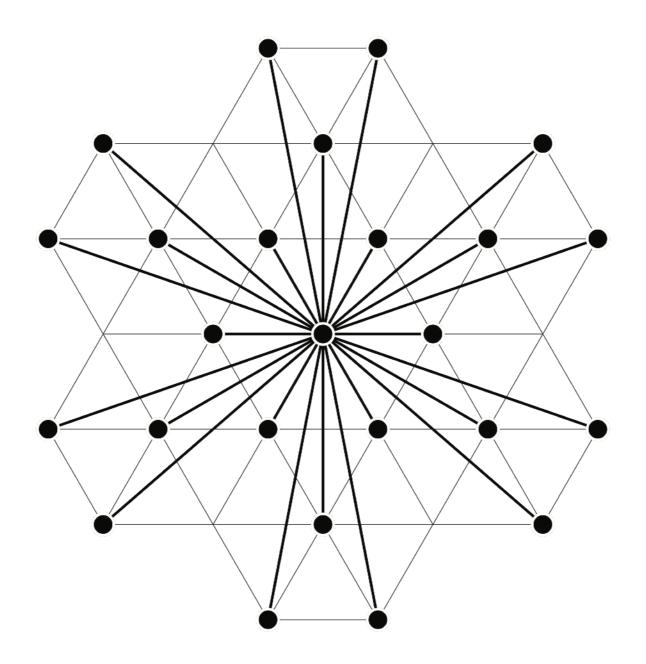
	List	Network
Dimension	One	Two
Max Connections	Two	Infinite
Metrics	Order	Distance / Proximity
Polarity	Vertical / Horizontal	Absent

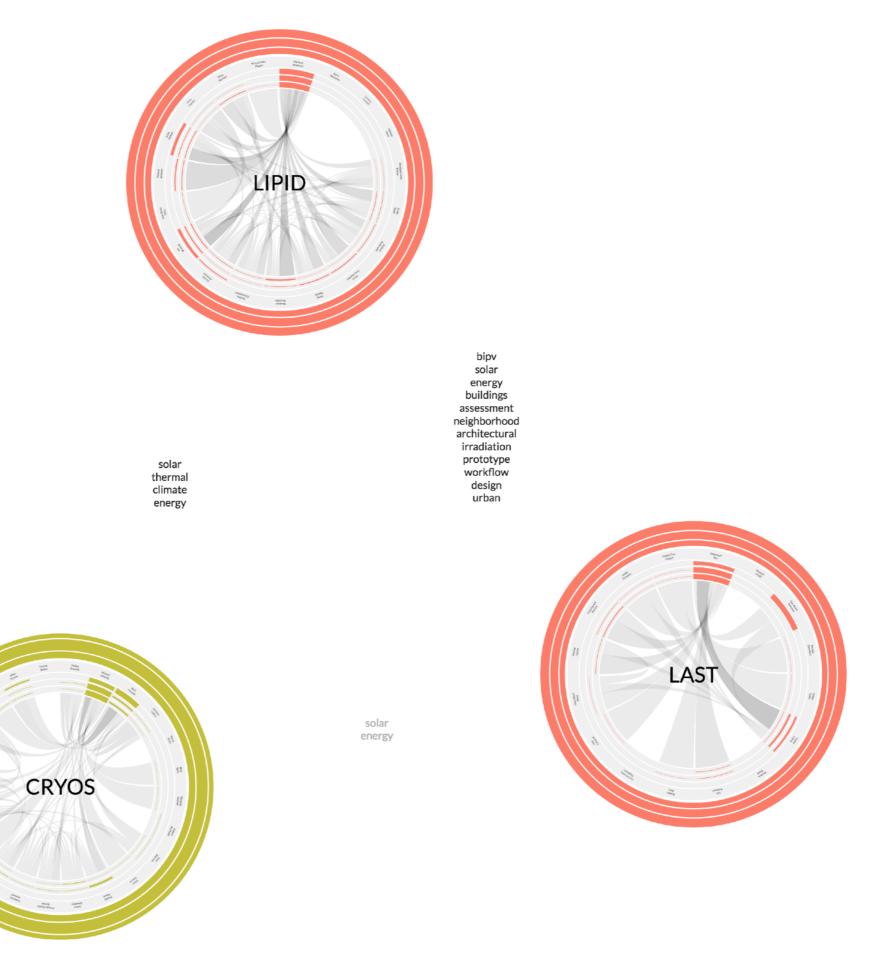






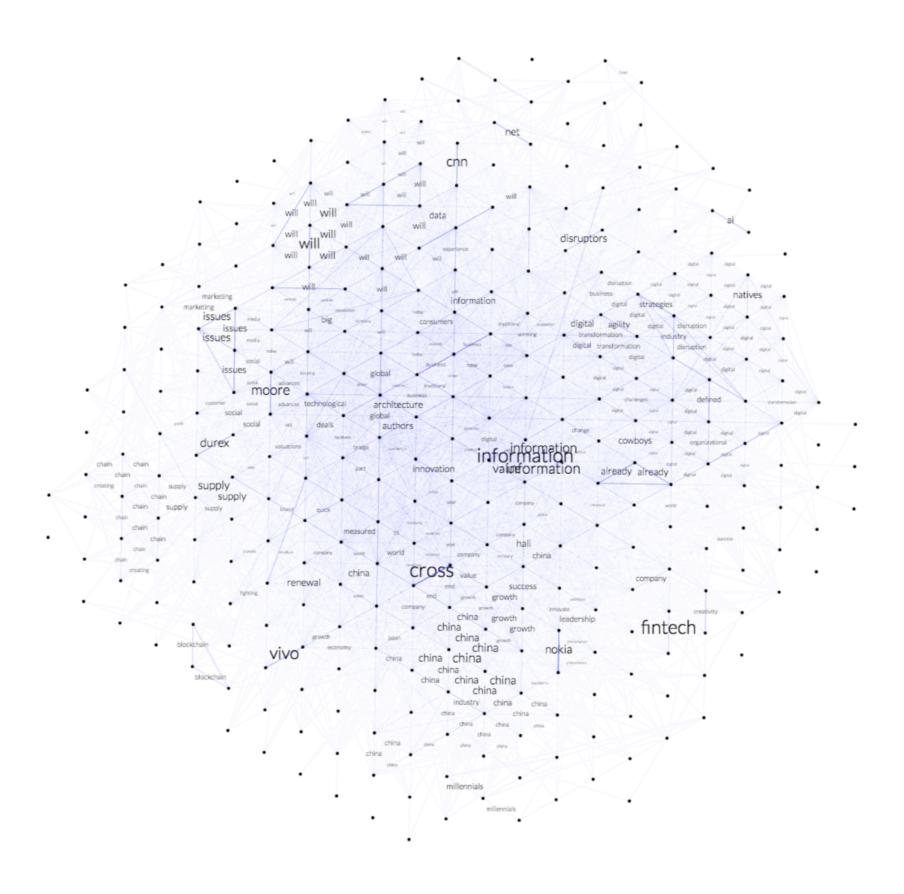






Mapping Affinities (Dario Rodighiero, 2018)

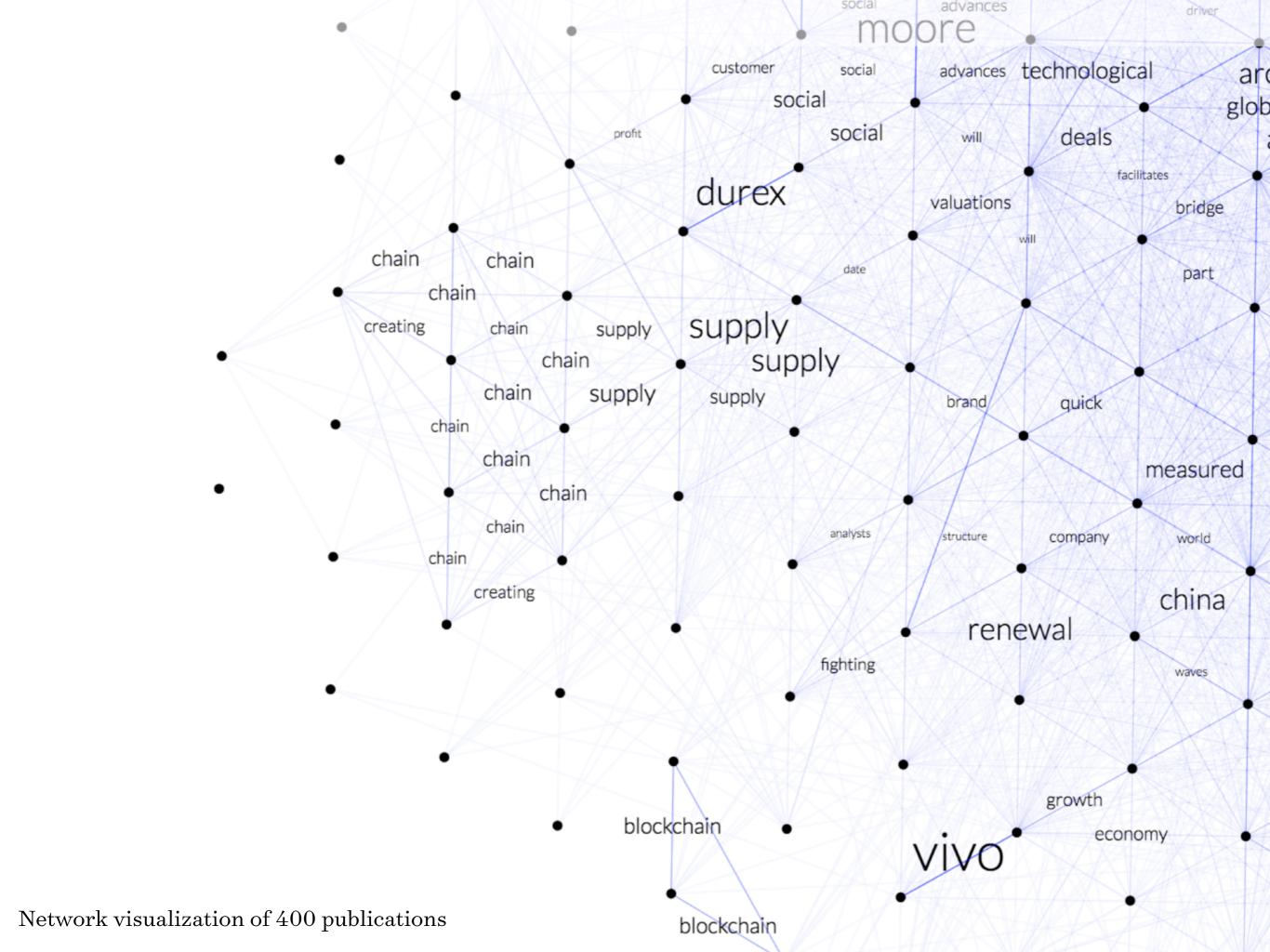
Network Visualization for Document Retrieval

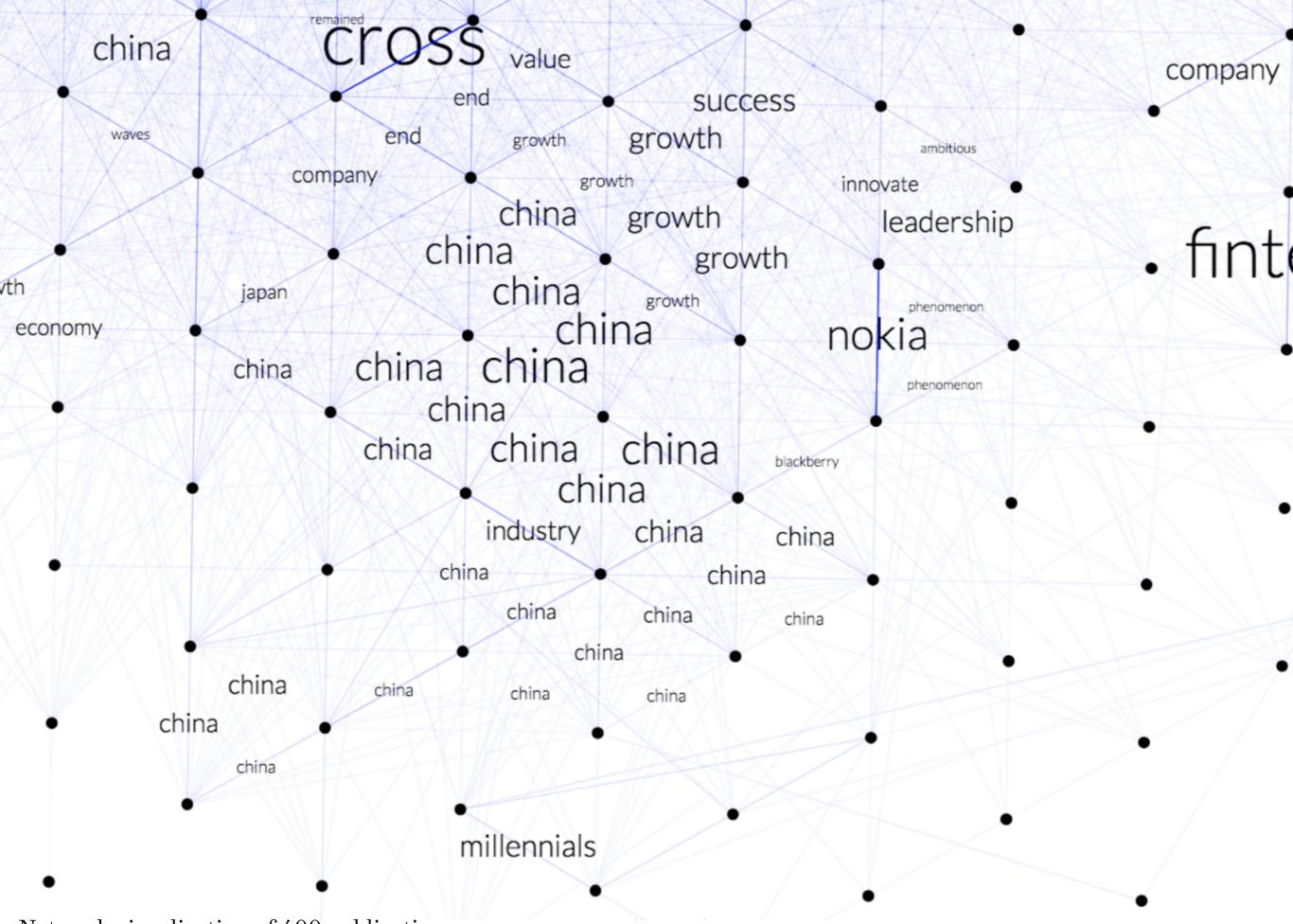


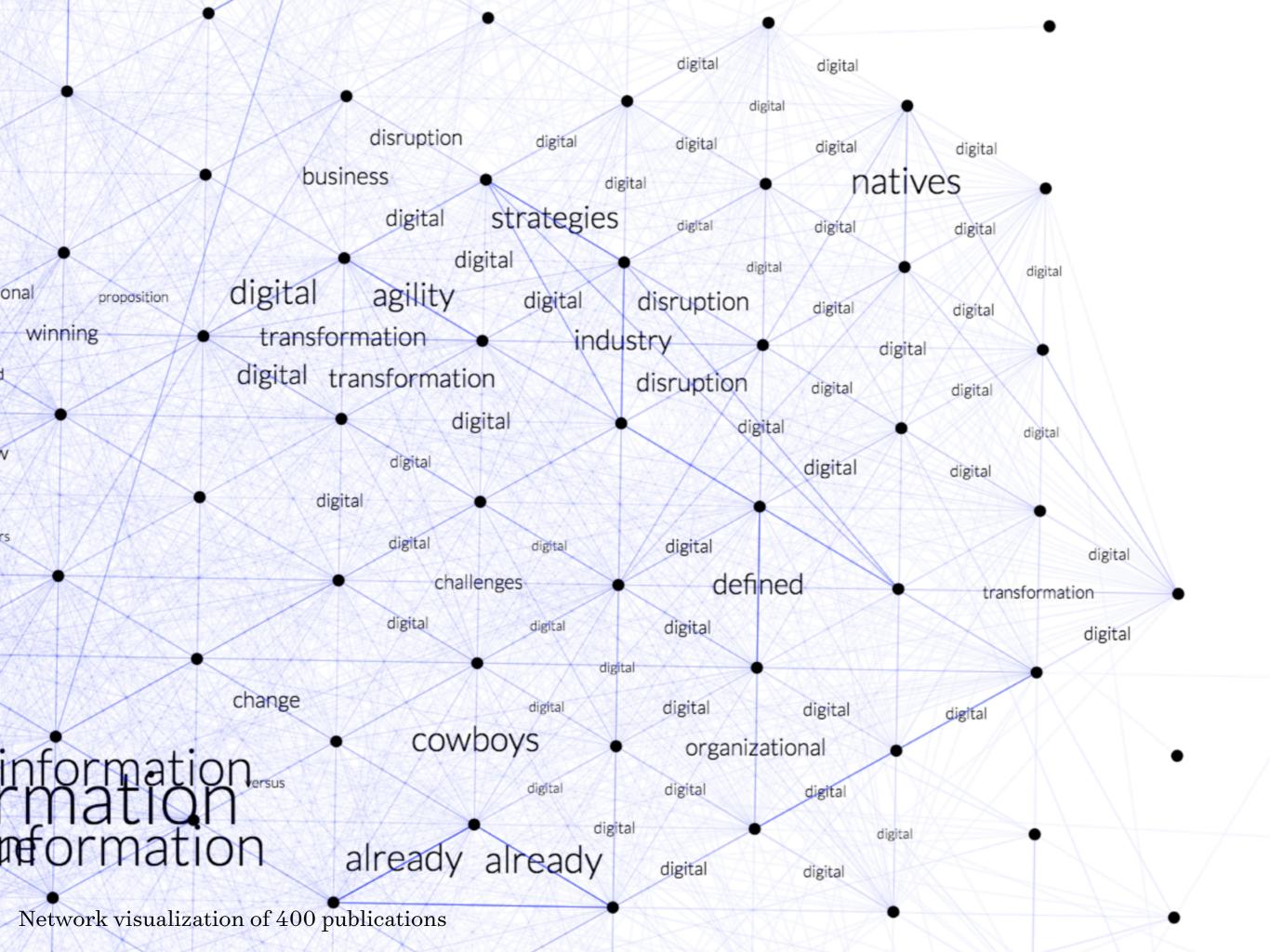
"In information retrieval, tf-idf or TFIDF, short for term frequency-inverse document frequency, is a numerical statistic that is intended to reflect how important a word is to a document in a collection or corpus."

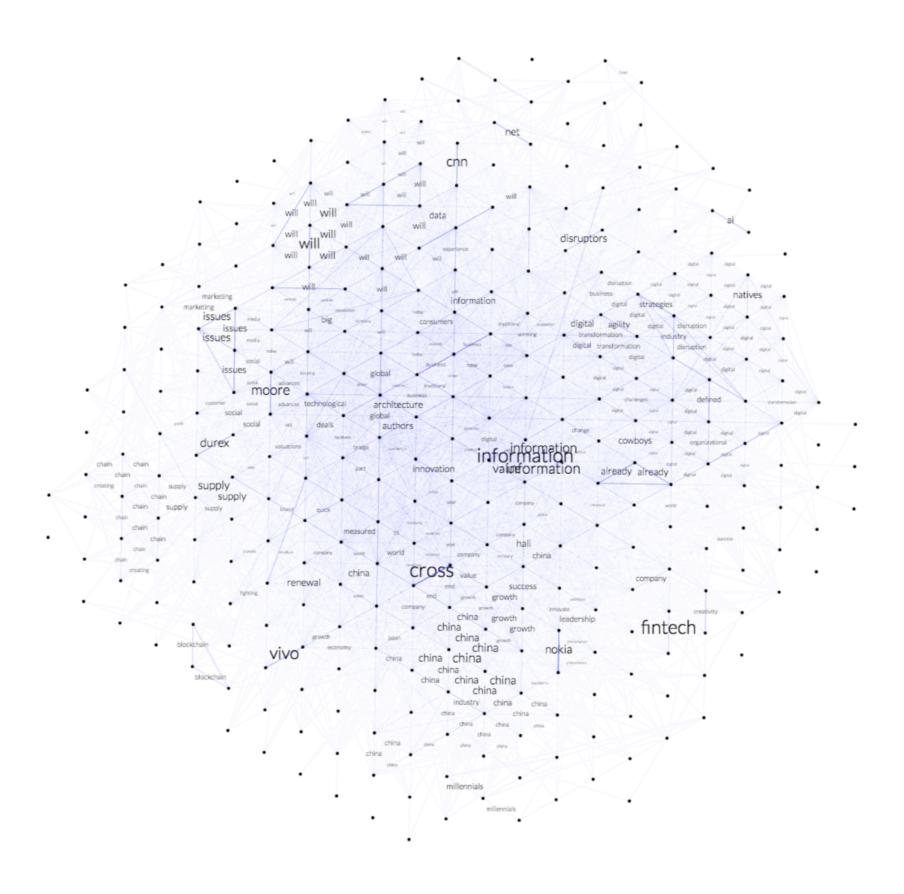
— Wikipedia 2018

```
{} docs.json ×
[ {
    "id": 53310,
    "authors": ["Bradley, Joseph", "Loucks, Jeff", "Macaulay, James", "Noronha, Andy", "Wade, Mi
    "title": "Digital vortex: How digital disruption is redefining industries",
    "description": "Digital business transformation is a journey to adopt and deploy digital ted
    "keywords": ["DIGITAL DISRUPTION", "DIGITAL BUSINESS TRANSFORMATION", "DIGITAL VORTEX"],
    "terms": {
        "digital": 13.493059996310013,
        "disruption": 12.672948161851869,
        "journey": 8.842000017916671,
        "models": 7.666426688112432,
        "technologies": 7.558292245571881,
        "inevitability": 6.030437921392435,
        "quantifiably": 6.030437921392435,
        "grasp": 6.030437921392435,
        "business": 5.884233501463874,
        "redefining": 5.624972813284271,
        "overturn": 5.624972813284271,
        "reshape": 5.624972813284271,
        "resulting": 5.624972813284271,
        "step": 5.624972813284271,
        "deploy": 5.624972813284271,
        "position": 5.33729074083249,
        "proposition": 5.33729074083249,
        "effect": 5.11414718951828,
        "force": 5.11414718951828
    "id": 54016,
```









#### **Conclusions**

The general idea is create a network that is able to semantically **fill the gap between nodes** that is usually unemployed.

Through terms we can create a **semantic background** that use semantic information to describe the **context** of each node.

This information is **complementary** to the visual network language that is based on points and lines.

### Next steps

Improve the quality of terms by using **N-grams** and a **controlled vocabulary** (Wikipedia).

Improve the interaction by using a **filter** function.

Visualize more terms for each proximity.

Test the visualization with more publications.

## Thanks

https://dariorodighiero.com