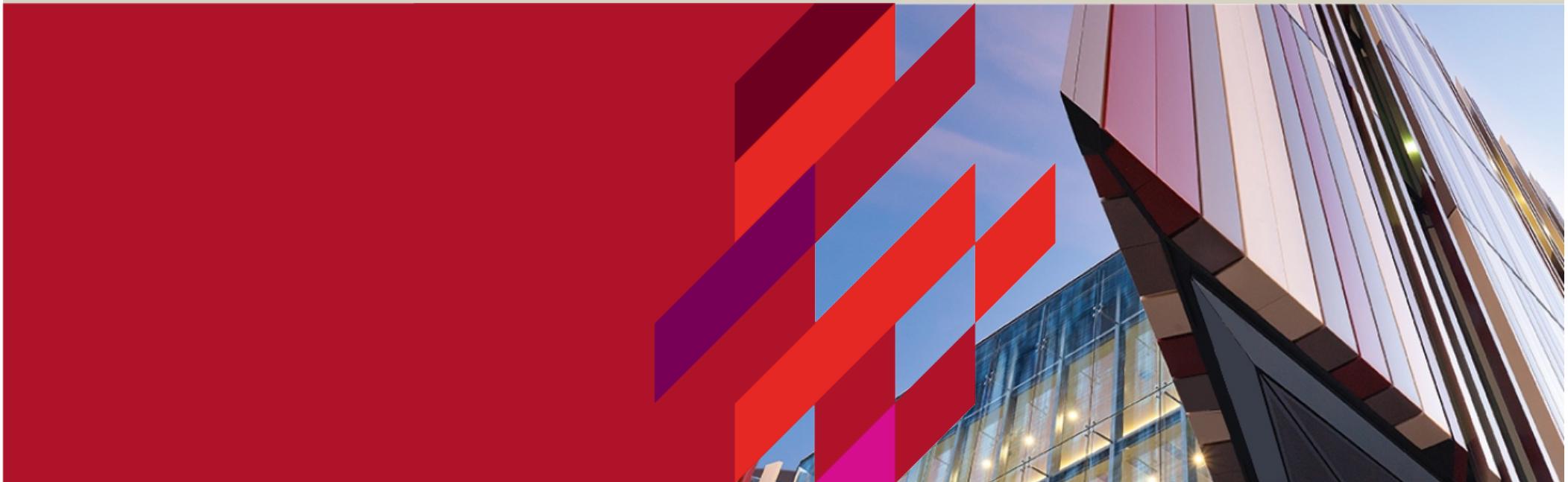




MACQUARIE  
University

# Automated Topic Extraction in Literature Reviews

Dr Mauricio Marrone





# We want to know how the research is evolving

---

- As researchers we want to know:
  - Topics are getting more/less attention
  - Changes in research themes
  - Gaps and opportunities
  - Overall - Monitor how the field evolves

Bibliometric  
mapping can  
help

- Maps the structure of knowledge in a field (K-structures):
  - Intellectual Structure – co-citation analysis
  - Social Structure – collaboration analysis
  - **Conceptual Structure – content analysis**
    - **Keywords, abstracts, titles.**

There are several tools that help analyse the content of papers

Tool
Bibliometric tools (like Bibliometrix, CiteSpace, etc)
Leximancer
LDA (Latent Dirichlet Allocation)

## Example of Issue (1 of 2)

- Similar concepts are often separated.
  - I.e.
    - Service Level Agreement
    - Service Level Agreements
    - SLA
    - SLAs
    - = 6 different concepts

## Example of Issue (2 of 2)

- Let us look at the term IT Service Management
  - IT? it?
  - Service?
  - Management?



## Enter Entity Linkers

- 1) Identify meaningful mentions in a text
- 2) link them with unambiguous identifiers drawn from a large catalogue or **knowledge base**, for example, Wikipedia.

Example:

Paris is the capital of France.

Paris? – Paris Hilton; Paris, New York; Paris, France ...

France? France River? Radio Station France?

Introducing  
the  
opportunity

# Example #1

Michael Jordan was a basketballer for the Bulls. Jordan played for the Dream team at the Olympics.

Topics identified:

- Michael Jordan
- Basketball
- Chicago Bulls
- 1992 United States men's Olympic basketball team
- Basketball at the summer Olympics

## Example #2

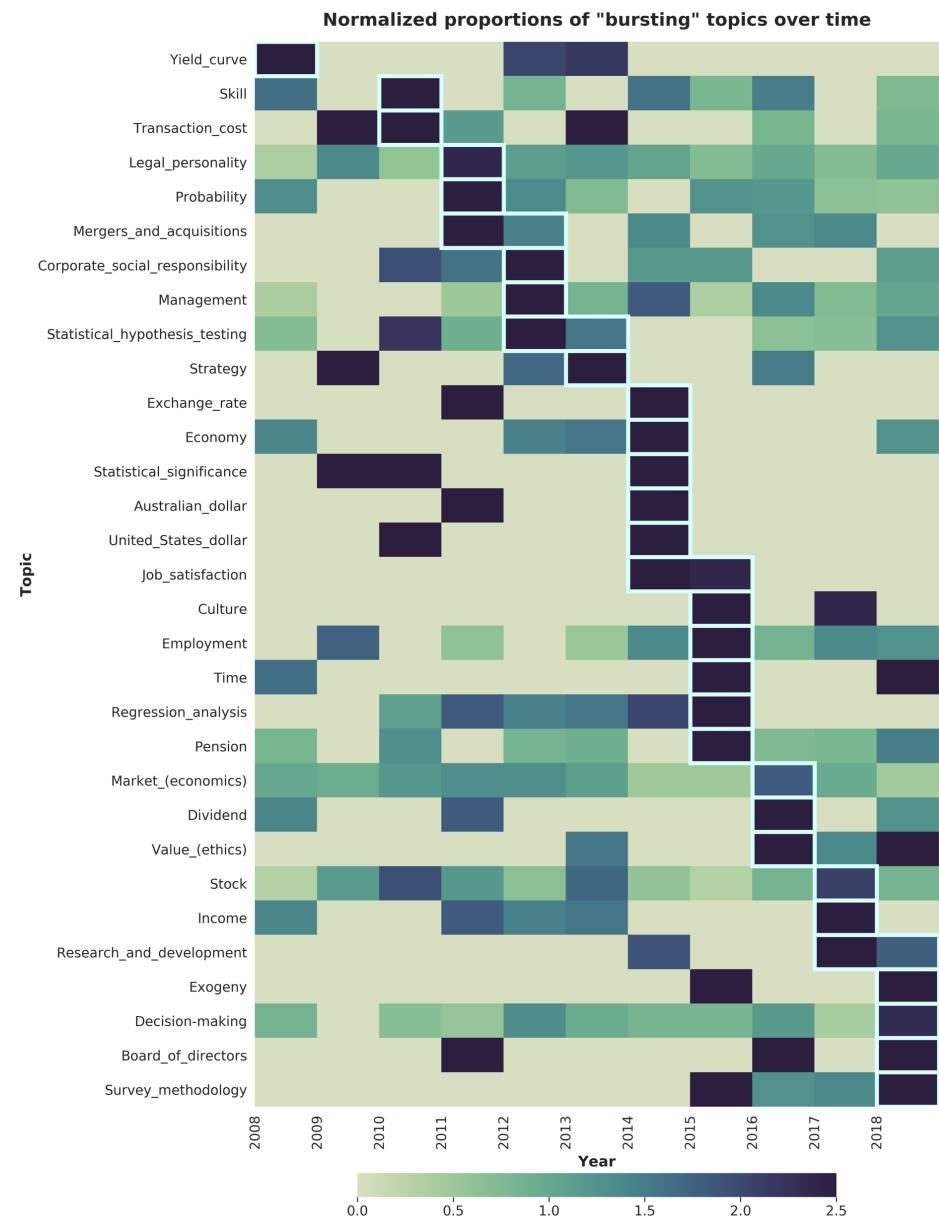
- *Homo sapiens* : *H. sapiens* ("wise man")
- **Binomial nomenclature**
  - binomial nomenclature
  - binary nomenclature
  - Latin name
  - Binomen name
  - binominal name
  - Scientific Name





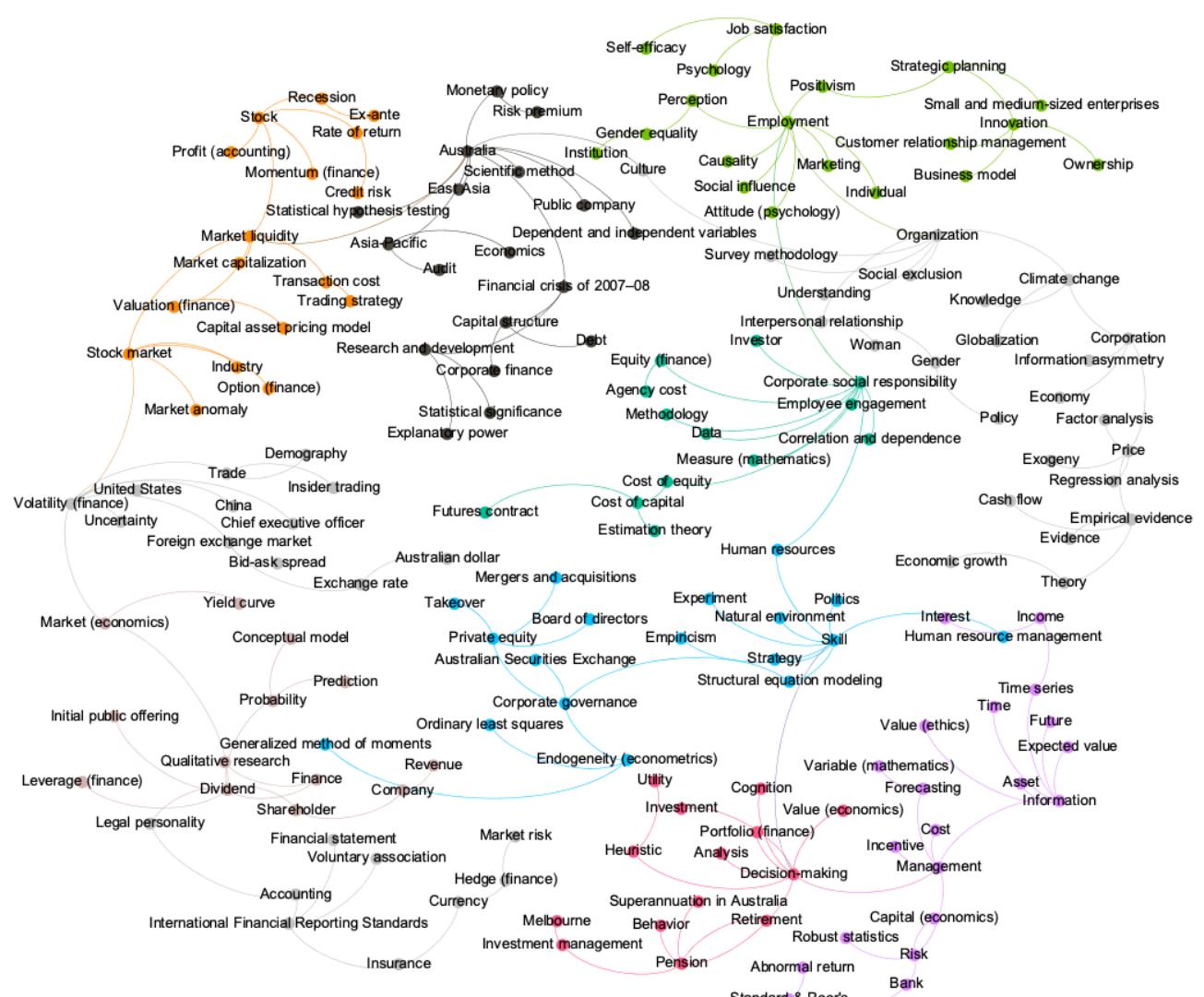
# Opportunity

This paper present ResGap, a tool which offers a way to overcome issues associated with homography and disambiguation.

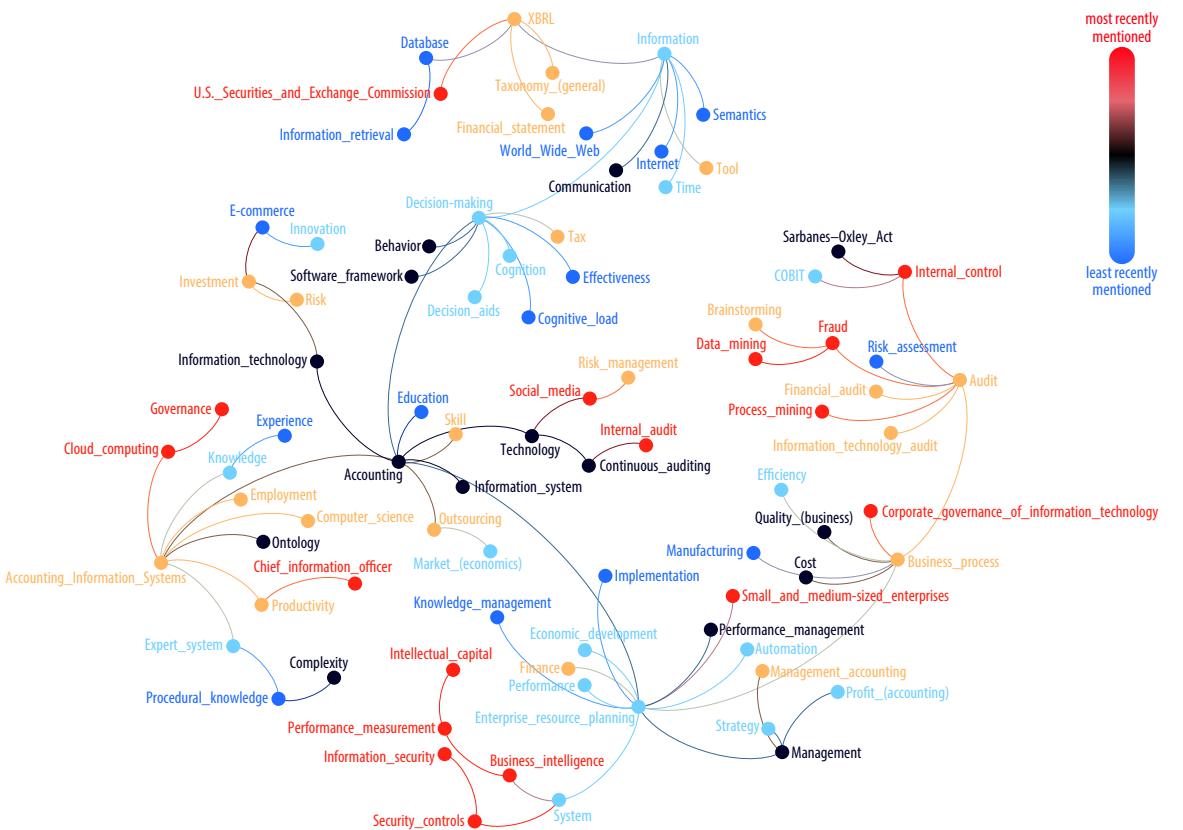


Example of Burstiness

# Example of Networks

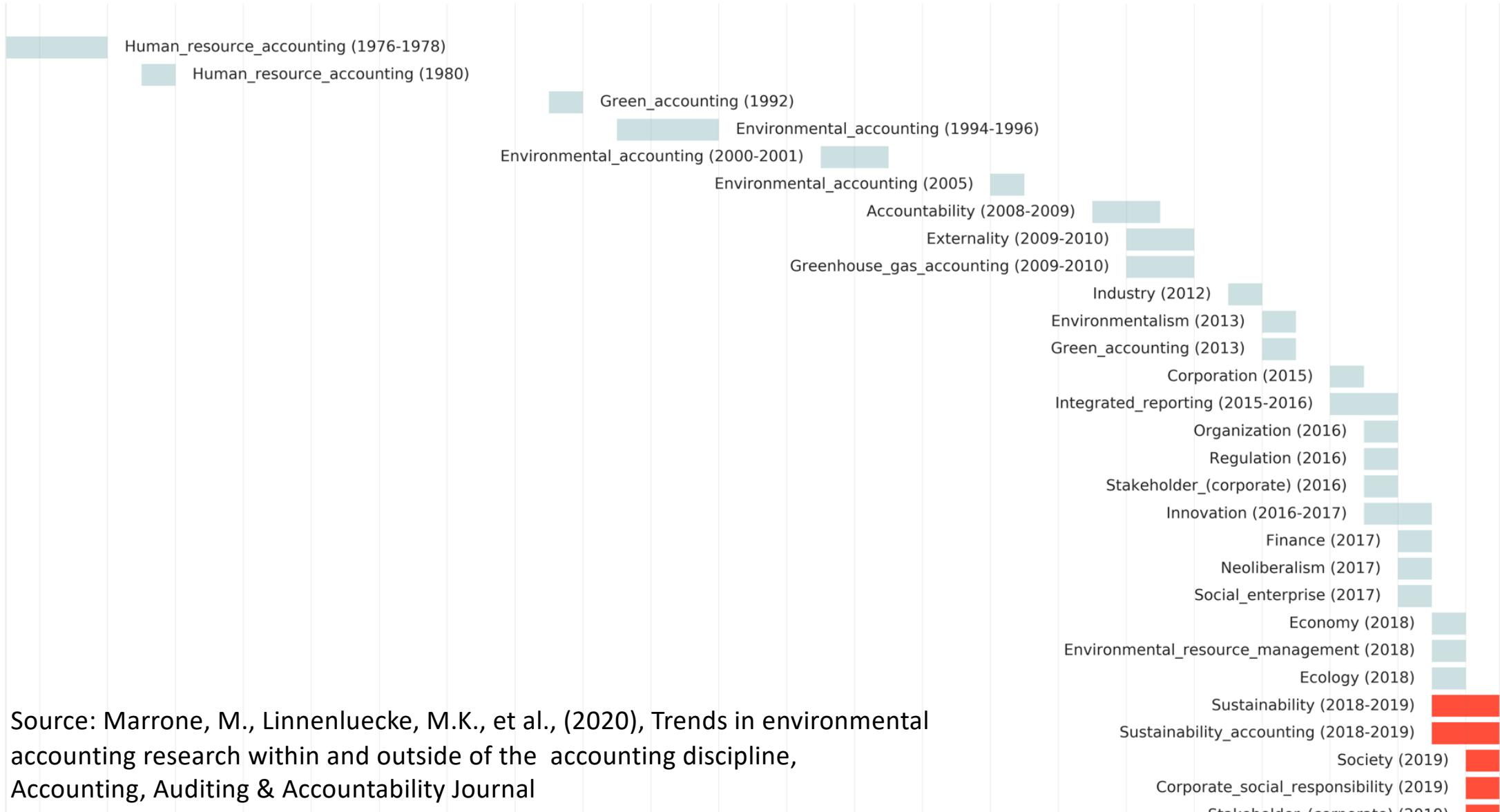


# Maximum Spanning Tree with Topical Focus Over Time

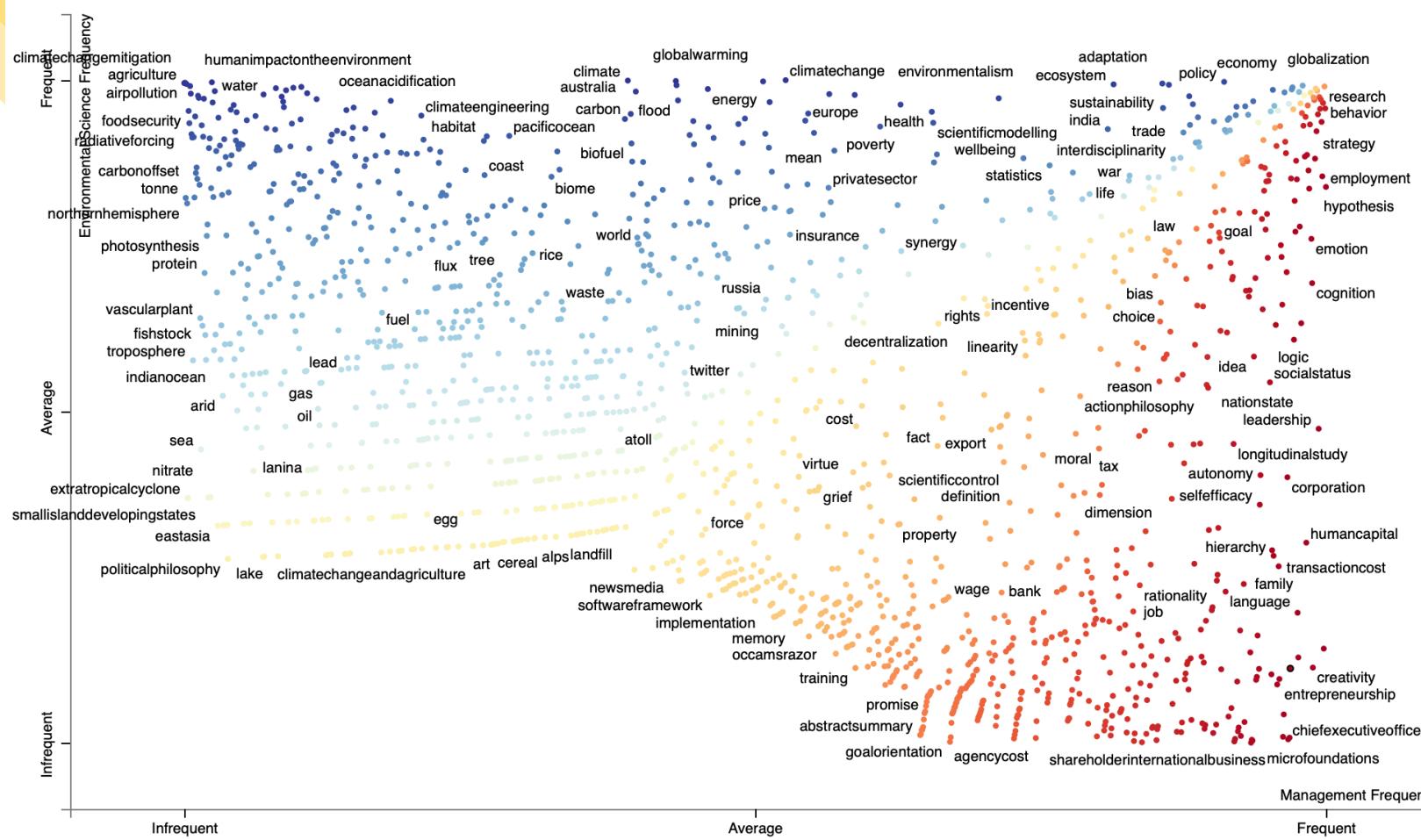


Source: Kumar, Marrone, et al., (2020), Twenty Years of the International Journal of Accounting Information Systems: A Bibliometric Analysis, International Journal of Accounting Information Systems

## Timeline of the top 30 "bursting" topics



Source: Marrone, M., Linnenluecke, M.K., et al., (2020), Trends in environmental accounting research within and outside of the accounting discipline, Accounting, Auditing & Accountability Journal



Top Environmental Science

climate  
greenhousegas  
temperature  
airpollution  
agriculture  
climatechange  
climatechangemitigation  
globalwarming  
humanimpactontheenvironment  
biodiversity  
carbon dioxide  
deforestation  
landuse  
earth

Top Management

multinationalcorporation  
creativity  
leadership  
humancapital  
performance  
cognition  
employment  
foreigndirectinvestment  
humanresourcemanagement  
corporatesocialresponsibility  
chiefexecutiveofficer  
emotion  
er interpersonalrelationship  
abusesupervision



MACQUARIE  
University

# Thank you

Mauricio.marrone@mq.edu.au

