

Daniel Richter, Philipp Baaden & Mohammad Saknini

Nordic Workshop on Bibliometrics and Research Policy 12.10.2023

Exploring the Thematic Landscape of the Circular Economy: A Comparative Study of Publication and Wikipedia Data

Exploring the thematic landscape of Circular Economy

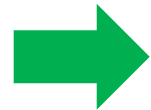
Our motivation:

Circular economy (CE) is one way to support transition towards a more sustainable future.

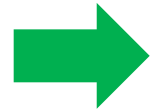
Need to identify technologies and innovations that support this transition.

Comparative analysis between established (Publication Data) and more unconventional (Wikipedia) Data sources

(Ena et al. 2016)



How can we contribute to this research with data driven / tech-mining methods?



What opportunities do Publication and Wikipedia Data offer to enable researchers, managers and policy makers to identify such technologies and trends?

Research aim:

Representing the thematic landscape of CE by identifying scientific knowledge fields in Publication & Wikipedia Data



Find thematic clusters in Publication and Wikipedia Data



Compare results and approaches between both resources



Find unique advantages and disadvantages of the datasets

The basis of our research

Definitions:

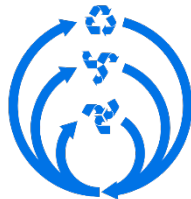
Research landscape consists of different scientific knowledge fields.

Scientific knowledge field is union of **scientific knowledge elements**, which are information embedded in a scientific article.

(Zhang et al. 2020)

Circular Economy:

- Minimizing waste generation and material inputs
- Maintaining value of products and resources
- Eco-design,
- Sharing, repairing, refurbishing,
- Recycling and reusing of products.



(Wijkman and Skanberg; 2015; Merli et al. 2018)

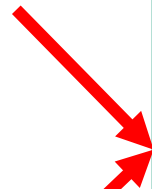


Source: [AI Art Generator - AI Image Generator API - Hotpot.ai](#)

Finding research streams and new scientific knowledge elements



Search query in
Dimensions database



Wikipedia Hyperlinks



Basic facts:

Query: "circular economy"

(Geissdoerfer et al., 2017)

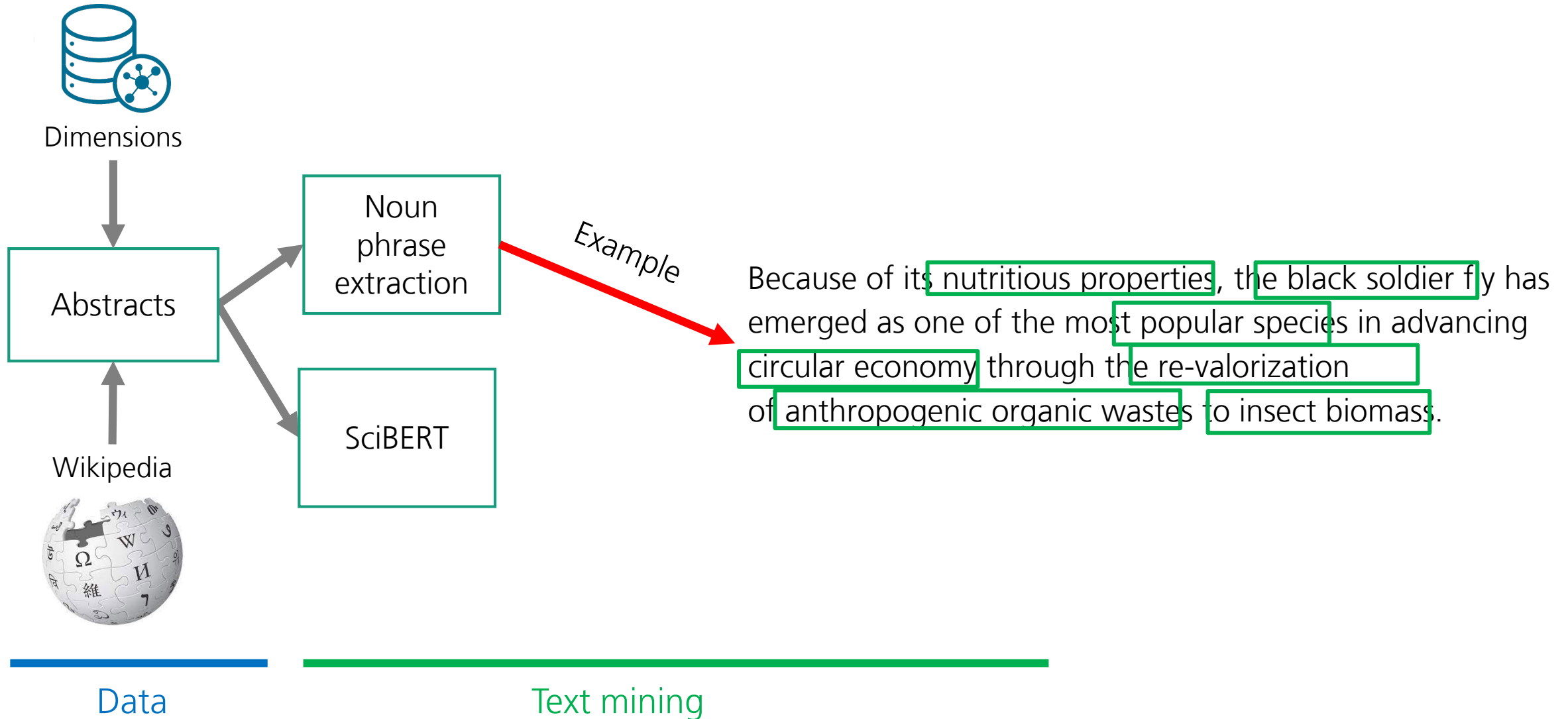
6,149 publications from 2023

Wikipedia article: „circular
economy“

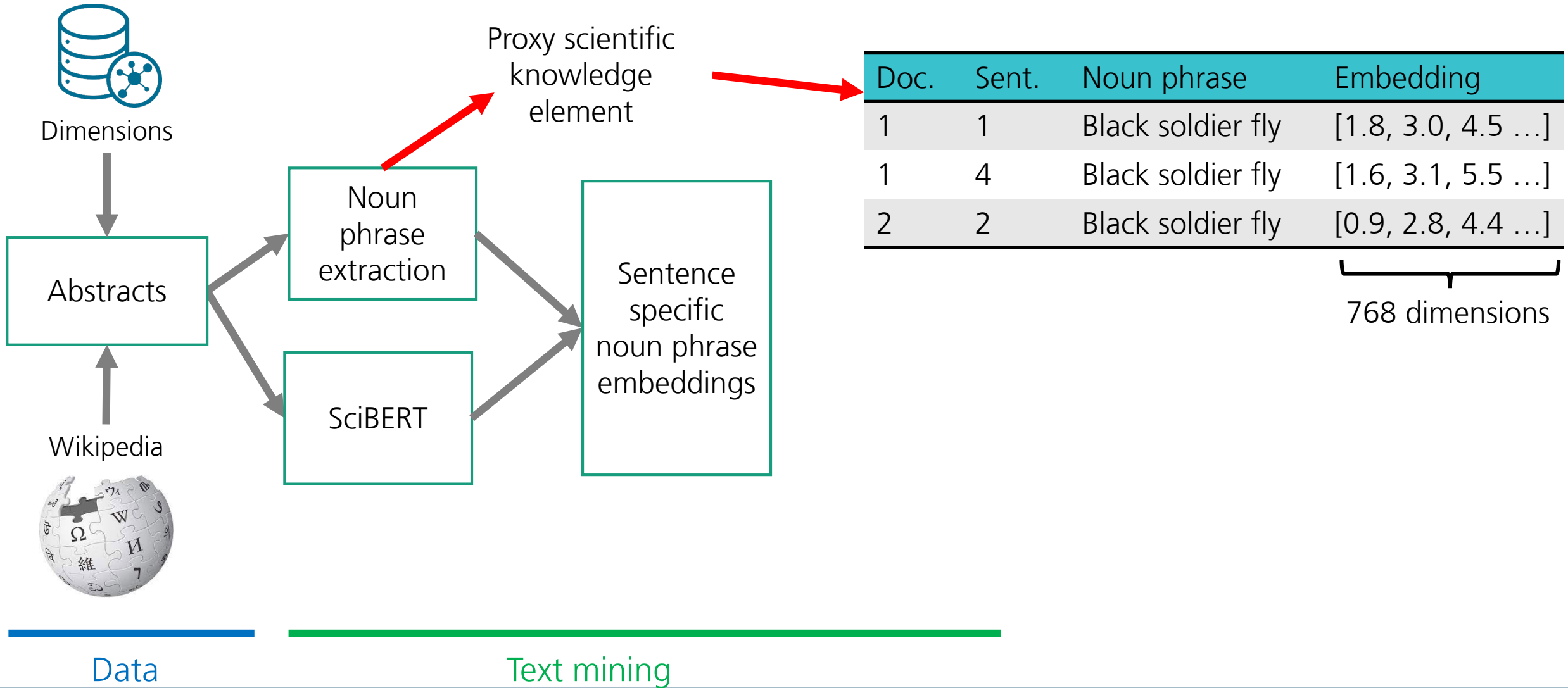
623 in- and outgoing hyperlinks to
connected articles

Data

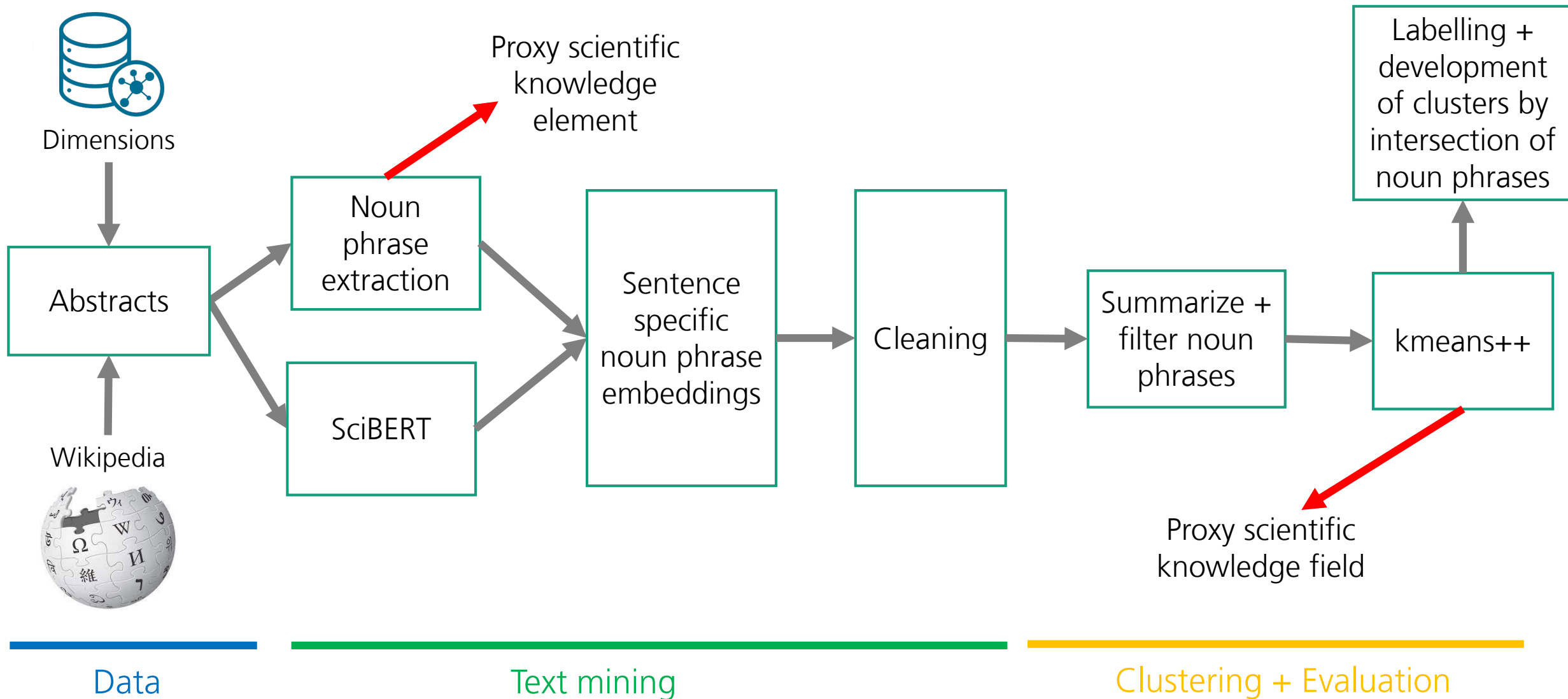
Finding research streams and new scientific knowledge elements



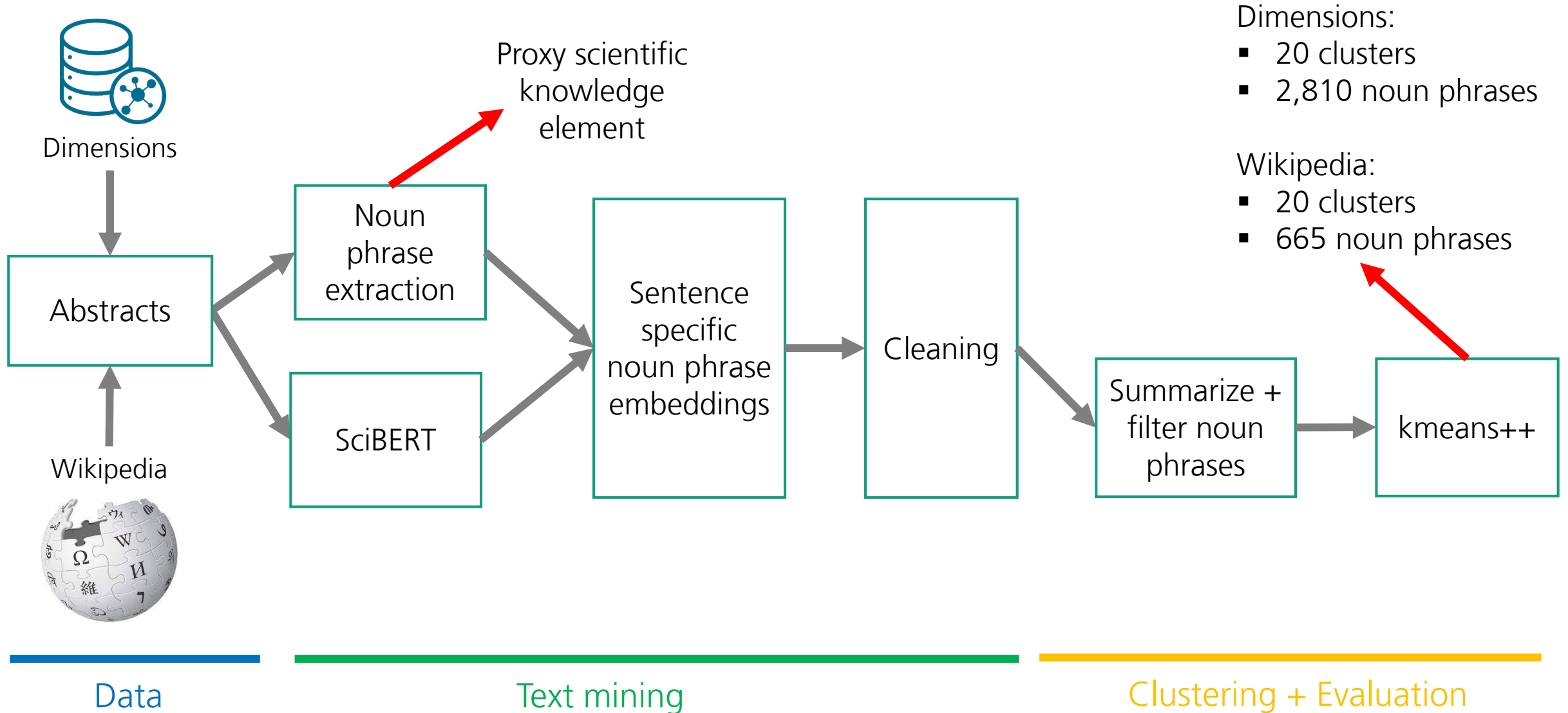
Finding research streams and new scientific knowledge elements



Finding research streams and new scientific knowledge elements



Finding research streams and new scientific knowledge elements



Clustering results

Cluster	Wikipedia Topic	Dimensions Topic
1	Carbon emission	Carbon emission
2	Wastes & Waste management	Wastes & Waste management
3	Sustainability & Recycling	Sustainability & Environment
4	Ecosystem & Lifecycles	Supply chains & Lifecycles
5	Natural resources & Resource management	Natural resources & Resource management
6	Share logistics & Recycle network	Share logistics & Recycle network
7	Globalisation & International Associations	Globalisation & Global demands
8	Business Management & Social sciences	Business Management & Social responsibility
9	Global demands	Urbanisation
10	Animal & human rights	Digitalisation
11	Green Politics & Policies	Challenges & Principles
12	Climate Change	Materials & Films
13	Energy production	Liquids & Indicators
14	Locations, Cities & States	Effects & Impacts
15	Costs & Markets	Organic materials
16	Environment & Technologies	Parameters, Approaches & Tools
17	-	Artificial Intelligence

Summary

- 4 clusters with identical topic ranges in both resources
- 4 clusters with very similar themes in Wikipedia and Dimensions
- 8 and 9 clusters with more individual topics between the two resources
- 4 and 3 clusters that were filled with stopwords or almost empty

Clustering results

Wikipedia Waste Cluster



Dimensions Waste Cluster

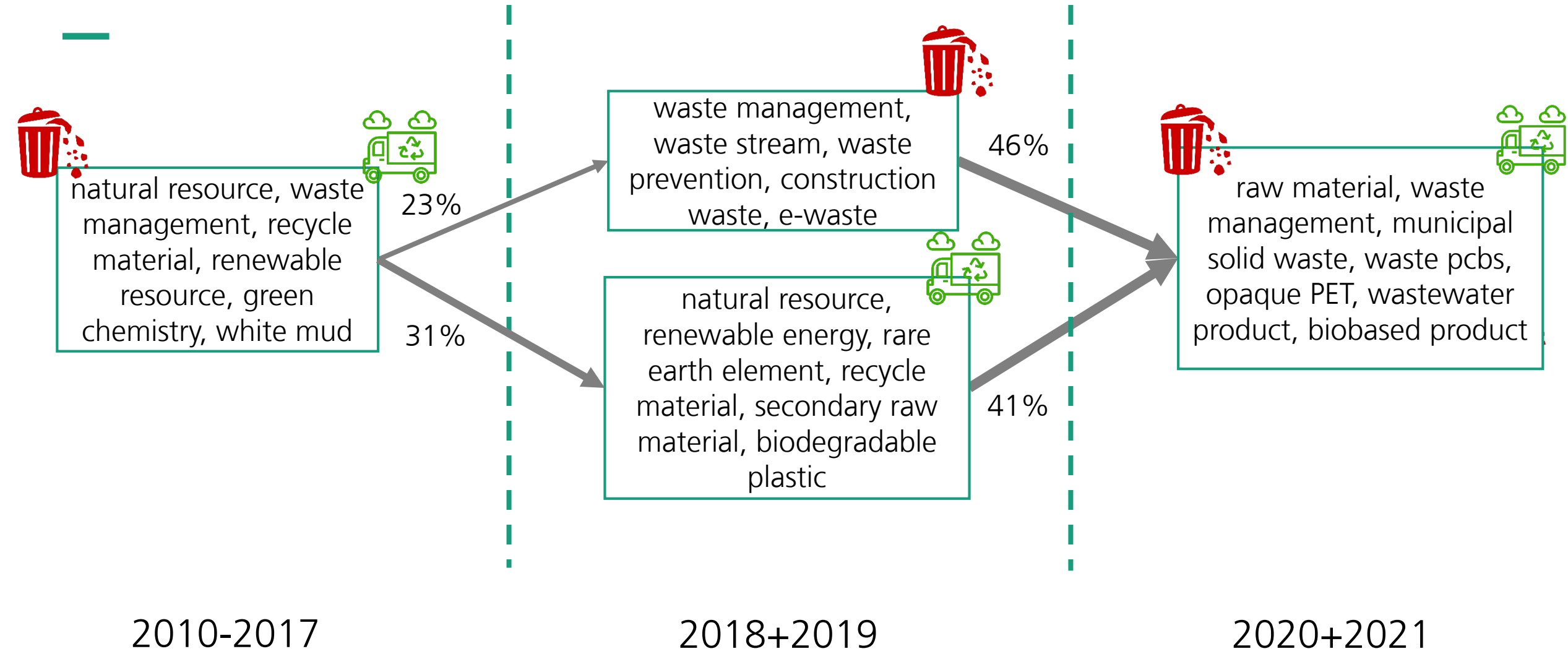


Evaluation of the two data resources

Which data resource is better?

- **Wikipedia and Publication data both allow for an analysis of the topic landscape of circular economy**
- **There are topics that can be found in both data resources but also individual clusters which leads to the assumption that a full analysis should include both approaches**
- **Wikipedia data includes a different perspective on the topic from publication data by offering more noun phrases including specific names of people, locations or associations**
- **It is difficult to evaluate the actual correctness or usefulness of the specific noun phrases in both cases**
- **Our approach for Wikipedia analysis does not yet allow an analysis over time to overview the development of a research landscape while it is possible with publication data thanks to the publication year**
- **The Wikipedia structure and concept of having one article per topic might be useful for a different approach of illustrating the topic landscape**

Unique for publication data: Finding research streams



Unique for publication data: Finding new knowledge elements

Cluster 2020+2021:

climate change, human health



New Knowledge elements:

carbon footprint, sustainability impact, social pressure, net zero emission

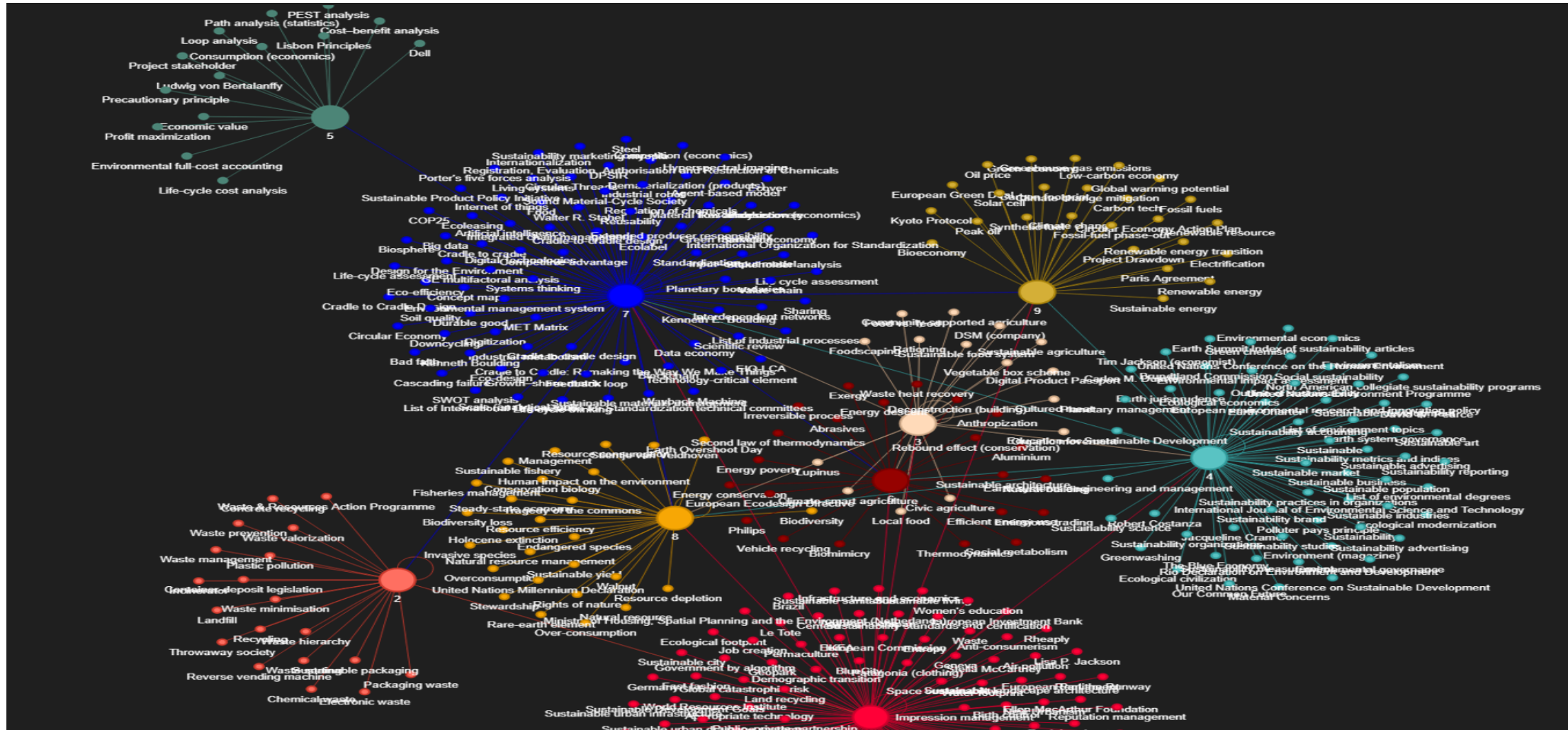
[Economic Assessment of Energy Consumption in Wastewater Treatment Plants: Applicability of Alternative Nature-Based Technologies in Portugal](#) /2022 [Details](#)

[...] In wastewater treatment systems, such an approach implies integrating efficient treatment technologies with aspects of the **circular economy**. [...] This technology, in addition to being financially advantageous, provides the ability to eliminate the **carbon footprint** and the valorisation of algae biomass, suggesting that this biotechnology is starting to position itself as a mandatory future solution in the wastewater treatment sector.

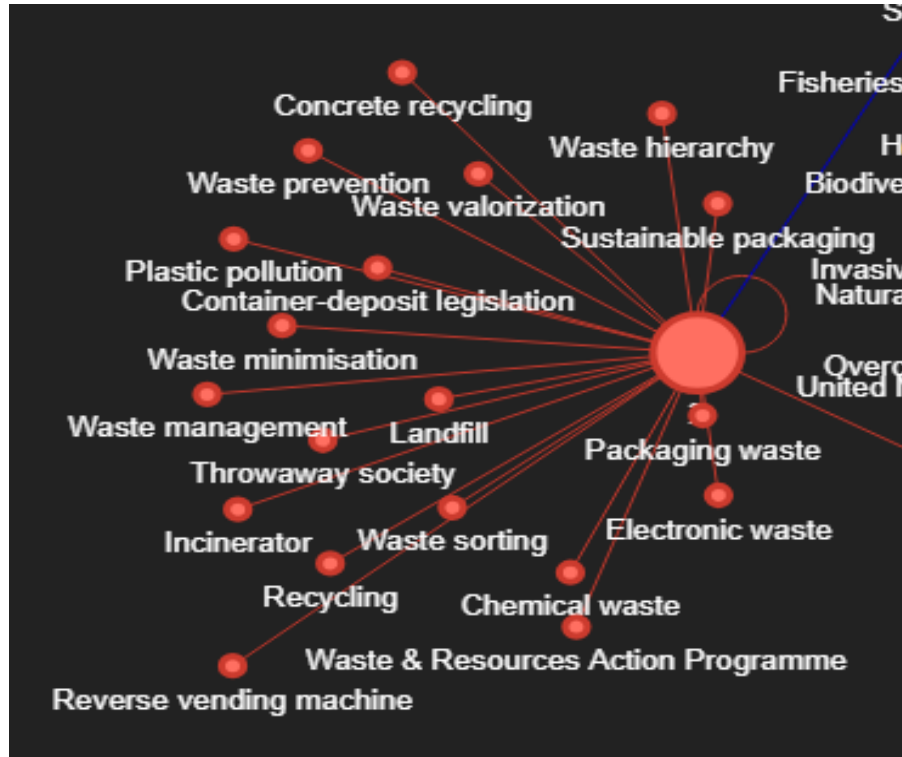
[Techno-economic feasibility and environmental sustainability of waste-to-energy in a **circular economy**: Sri Lanka case study](#) /2022 [Details](#)

[...] The study presented in the paper has examined WtE in a **circular economy** as a solution that can have economic, financial, social, and environmental co-benefits through efficient use of natural resources, reduced emissions, and fostering innovation. The case study involving Sri Lanka has shown that similar countries have the potential to implement WtE projects coupled with appropriate **circular economy** elements with adequate financial and economic returns. [...] This pioneering case study can be used as a basis for immediate action to solve waste management issues within a **circular economy** helping similar economies in the Asia Pacific region in their efforts to achieve **net zero emission** target by the middle of this century.

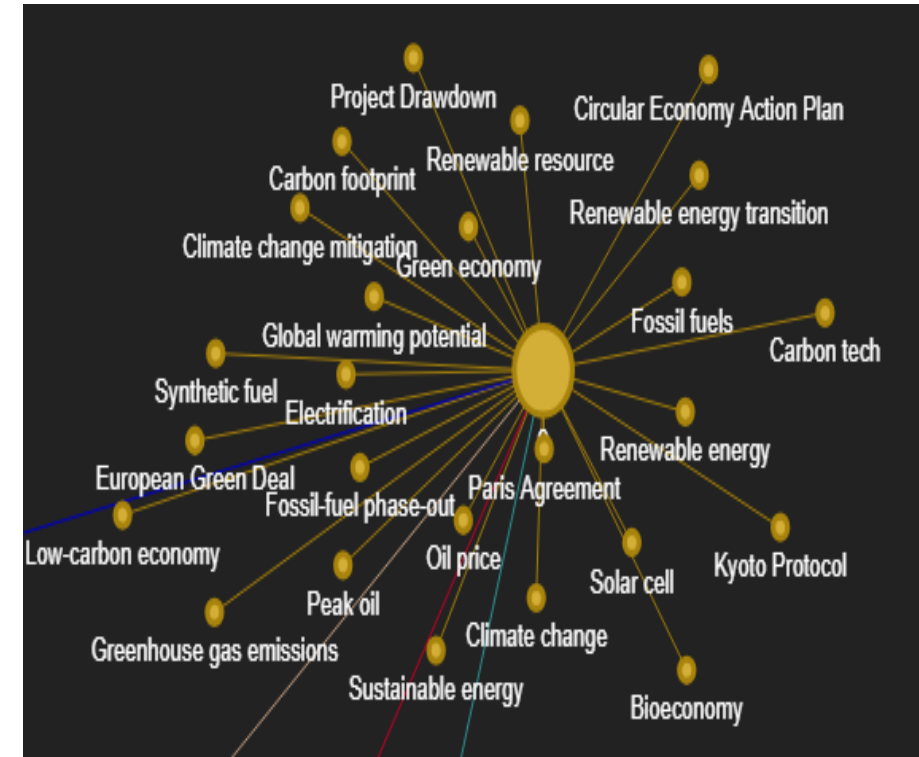
Unique for Wikipedia data: Hyperlink network illustration



Unique for Wikipedia data: Hyperlink network illustration



Wikipedia articles about waste and waste management



Wikipedia articles about energy production and emissions

Conclusions, Limitations and Future work

Conclusions:

- Both data resources allow to investigate clusters and the thematic lands 🎯
- The clusters show similarities but also differences between the data resources 🎯
- The approach illustrates unique strengths and opportunities but also weaknesses of Wikipedia and Publication Data for bibliometric analysis 🎯

Limitations:

- Design decisions e.g. cluster labelling, selection of hyperlinks,...
- Need for extended and research field specific stopword list
- No automatic analysis of the clusters possible at this point

Future work:

- Improving information extraction
- Expert evaluation
- Investigate the inclusion of the second layer of hyperlinks in Wikipedia
- Find a qualitative method to measure the quality difference between Wikipedia and Publication data



Daniel Richter
daniel.richter@int.fraunhofer.de

Fraunhofer Institute for
Technological Trend Analysis (INT)

Appelsgarten 2, 53879 Euskirchen