

Year

2020

### Author(s)

Haiyun Xu and Jos Winnink and Zenghui Yue and Ziqiang Liu and Guoting Yuan

### Title

Topic-linked innovation paths in science and technology

#### Venue

Journal of Informetrics

# Topic labeling

Partially automated

### **Focus**

Secondary

# Type of contribution

Novel approach

## Underlying technique

Manual labeling assisted by automatic summarisation

# Topic labeling parameters

١

## Label generation

LexPageRank (Erkan & Rade, 2004) method is used to obtain an automatic summary for each topic.

#### 1.1 Scientific topics

**S0:** Herpes simplex virus as a vector for cancer immunotherapy, studies of expression of vaccine proteins in plant cells, and engineering of bacteria in yeast. Herpes simplex virus can replicate in cancer cells but cannot replicate in normal cells, and herpes simplex virus can be used as an anticancer agent by recombining genes that can help the tumor to induce acquired immunity. A focus of new vaccine technology research, plants have a complete eukaryotic expression system, can be safely consumed, and have no pathogen pollution. On the other hand, direct manipulation of the bacterial genome is often difficult to achieve. The genome of *M. mycoides* was cloned into yeast artificial chromosomes, genetically engineered, and then transplanted into *M. capricolum*.

### 1.2 Technological topics

T0: Construction of a nucleic acid vaccine. The genes encoding the pathogen antigen and the plasmid DNA as an expression vector are included in the treatment of herpes simplex virus, B lymphocyte-derived tumor, and the like from the disease type.

Under the guidance of domain experts, the content of 15 scientific topics and 15 technical topics with clear connotations was obtained.

#### Motivation

\

## Topic modeling

IDA

### Topic modeling parameters

Nr of topics (k): 15

# Nr. of topics

15 scientific topics, and 15 technical topics

### Label

### Label selection

\

### Label quality evaluation

\

### Assessors

\

### **Domain**

Domain (paper): Scientific innovation

Domain (corpus): Medical (GEV)

#### Problem statement

This study focuses on connected topics utilizing bibliometric analysis, thereby exploring the identification method for innovation paths based on the linkage of scientific and technological topics.

The internal mechanism of knowledge dissemination and the relationship between science and technology are revealed and described in detail by measuring the linkage of knowledge units.

For practical bibliometric analyses, research papers and patent literature were used to characterize scientific research and technological research to reveal the innovation path for the interaction of science and technology quantitatively, automatically, and visually.

### Corpus

Origin: Clarivate Analysis (Web of Science database (WoS), Derwent Innovations Index database (Derwent))

Nr. of documents: 8196 (4146 records for scholarly publications and 4050 records for patent publications)

#### Details:

• Research papers and patent literature were used to characterise scientific research

and technology research

- In this study, the GEV field was selected as our experimental field. GEV is a research area that is expected to be the source of new innovative vaccines.
- Papers and patent documents are collected in the GEV field for a period as analysis data sets.
- The Web of Science database (WoS) is selected to search scientific papers.
- The Derwent Innovations Index database (Derwent) is used to search patent documents.
- Data collection was performed on January 6, 2018, and the publication year is up to 2017.

### **Document**

title, abstract, author, and cited journal of research or patent paper in the GEV domain.

### Pre-processing

Keyword phrases from the patents and the abstracts of the articles are extracted and added to the subject field of patents and papers.

@article{xu\_2020\_topic\_linked\_innovation\_paths\_in\_science\_and\_technology, abstract = {In the modern world, science and technology jointly determine the evolutionary path of scientific innovation, with an increasingly close relationship between them. Therefore, it is important to study the identification method of the innovation path, based on the linkage of topics in science and technology. This study focuses on connected topics utilizing bibliometric analysis, thereby exploring the identification method for innovation paths based on the linkage of scientific and technological topics. The internal mechanism of knowledge dissemination and the relationship between science and technology are revealed and described in detail by measuring the linkage of knowledge units. For practical bibliometric analyses, research papers and patent literature were used to characterize scientific research and technological research to reveal the innovation path for the interaction of science and technology quantitatively, automatically, and visually. Experimental study shows that analysis of the topic-linked path of science and technology, along with the integration of multi-relationships, can effectively

```
identify important science- and technology-related topics in a field in the
evolution process, and help grasp the key points of basic research and applied
research.},
  author = {Haiyun Xu and Jos Winnink and Zenghui Yue and Ziqiang Liu and
Guoting Yuan},
  date-added = \{2023-03-14 \ 20:41:48 \ +0100\},
  date-modified = \{2023-03-14\ 20:41:48\ +0100\},
  doi = {https://doi.org/10.1016/j.joi.2020.101014},
  issn = \{1751-1577\},
  journal = {Journal of Informetrics},
  keywords = {Evolution path, Scientific innovation, Science-technology
linkage, Knowledge dissemination},
  number = \{2\},
  pages = \{101014\},
  title = {Topic-linked innovation paths in science and technology},
  url = {https://www.sciencedirect.com/science/article/pii/S175115771930210X},
  volume = \{14\},
  year = {2020},
  bdsk-url-1 = {https://www.sciencedirect.com/science/article/pii/
S175115771930210X},
  bdsk-url-2 = {https://doi.org/10.1016/j.joi.2020.101014}}
```

#### #Thesis/Papers/Initia