

SciTo trends: visualising scientific topic trends

Serafeim Chatzopoulos^{1,2}, Panagiotis Deligiannis¹, Thanasis Vergoulis², Ilias Kanellos^{2,3}, Christos Tryfonopoulos¹, Theodore Dalamagas²

¹Univ. of the Peloponnese, Dep. of Informatics & Tel/tions, Tripoli, Greece

²IMSI, "Athena" Research & Innovation Center, Athens, Greece

³NTUA, School of Electrical & Computer Engineering, Athens, Greece

Data collection and processing

Article citation data

- ▷ OpenCitations COCI dataset¹
- ▷ ~ **450M** citations for > **45M** articles

Article impact scores

- ▷ Citation counts (overall impact)
- ▷ RAM [1] scores (short-term impact)
- ▷ Both impact scores are gathered by BIP! Finder [2]

Article abstracts

- ▷ ~ **12M** abstracts were collected from Open Academic Graph² [3, 4] and Crossref API³

- ▷ An **LDA** [5] model was trained using the gensim⁴ library

¹ <https://opencitations.net/download>

² <https://www.openacademic.ai/oag/>

³ <https://www.crossref.org/services/metadata-delivery/rest-api/>

⁴ <https://radimrehurek.com/gensim/>

SciTo's search interface

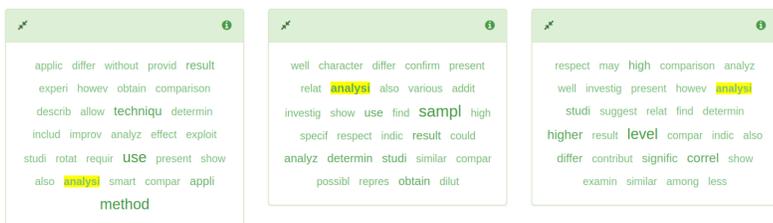
- ▷ Keyword-based search interface based on Apache Solr¹ that facilitates scientific topics exploration



analysis Find

20 results (1 pages)

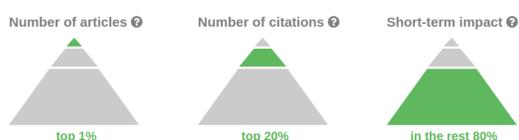
Click on entries for comparison



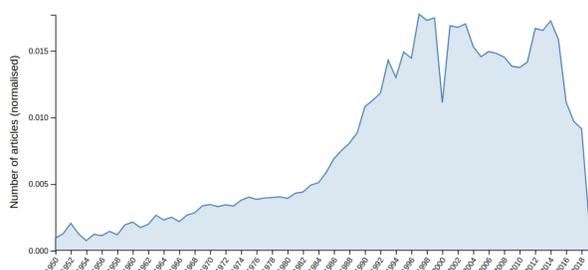
¹ <https://lucene.apache.org/solr/>

Visualisations for topic details

- ▷ Popularity trends for a specific topic
- ▷ Intuitive infographics based on 3 indicators:
 - ▷ Number of topic-related articles published
 - ▷ Number of citations attracted by topic-related articles
 - ▷ Average short-term impact for topic-related articles published
- ▷ Two different types of infographics:
 - ▷ Pyramid infographic
 - ▷ Shows if topic is at the top 1% or 20%
 - ▷ Trend infographic
 - ▷ Displays evolution of the 3 popularity indicators



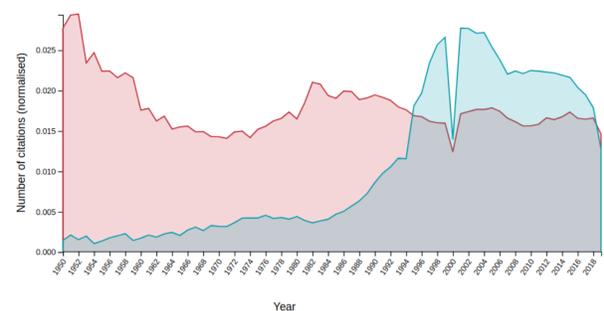
Number of articles per year Citations per year Short-term impact per year



Visualisations for topic comparison

- ▷ 2 or more topics can be selected for comparison
- ▷ The comparison view contains the Trend infographic for all topics under comparison
- ▷ Comparison scenario:
 - ▷ A researcher explores topics related to the keyword "gene". She finds a rather popular topic containing the terms "mrna", "rna" and "transcript". (blue time-series)
 - ▷ Then, she wants to compare this topic with the research field studying drug effects. She identifies a topic from life sciences containing the terms "drug", "treatment" and "effect" (red time-series).
 - ▷ In the comparison view, the Trend infographic reveals that although the "drug"-related topic was traditionally more popular, after **1995** the "gene"-related topic started to become equal or more popular (depending on the indicator used)

Number of articles per year Citations per year Short-term impact per year



Normalisation per year

References

- [1] Rumi Ghosh, Tsung-Ting Kuo, Chun-Nan Hsu, Shou-De Lin, and Kristina Lerman. Time-aware Ranking in Dynamic Citation Networks. In *IEEE ICDMW*, pages 373–380. IEEE, 2011.
- [2] Thanasis Vergoulis, Serafeim Chatzopoulos, Ilias Kanellos, Panagiotis Deligiannis, Christos Tryfonopoulos, and Theodore Dalamagas. Bip! Finder: Facilitating scientific literature search by exploiting impact-based ranking. *CIKM*, 2019 (to appear).
- [3] Arnab Sinha, Zhihong Shen, Yang Song, Hao Ma, Darrin Eide, Bo-june Paul Hsu, and Kuansan Wang. An Overview of Microsoft Academic Service (MAS) and Applications. In *WWW*, pages 243–246, 2015.
- [4] Jie Tang, Jing Zhang, Limin Yao, Juanzi Li, Li Zhang, and Zhong Su. ArnetMiner: Extraction and Mining of Academic Social Networks. In *ACM SIGKDD*, pages 990–998, 2008.
- [5] David M. Blei, Andrew Y. Ng, and Michael I. Jordan. Latent Dirichlet Allocation. *JMLR*, 2003.

Funding: We acknowledge support of this work by the project "Moving from Big Data Management to Data Science" (MIS 5002437/3) which is implemented under the Action "Reinforcement of the Research and Innovation Infrastructure", funded by the Operational Programme "Competitiveness, Entrepreneurship and Innovation" (NSRF 2014-2020) and co-financed by Greece and the European Union (European Regional Development Fund).



Co-financed by Greece and the European Union