Topic modeling

In Research.fi portal's topic modeling unsupervised machine learning methods (topic modeling) were used to algorithmically cluster nearly 7 000 projects into topics based on their title, short project description, possible keywords, and fields of science. Several unsupervised topic modeling methods (*contextualized topic model, top2vec, hierarchical stochastic block model, BERTopic*) were compared to identify topics.

After comparisons, 92 topics were selected. Each project is classified to a single, most likely topic. Each topic is algorithmically labeled with their most significant keywords. As the keywords describe the common characteristics of each topic, they may not be fully descriptive for every project in the topic. Furthermore, all projects may not have been clustered correctly into their most appropriate topic. Therefore, the topic modeling results are not suitable for accurately determining the sizes of individual topics.

New projects are classified to the existing topics daily. When a considerable amount of new data is available, the topic modeling will be performed again to see whether entirely new topics should be added. In this case, the number of topics and topic keywords may change, and individual projects may move between topics.

More information on topic modeling methods

Contextualized topic models

- Bianchi, F., Terragni, S., Hovy, D., Nozza, D., & Fersini, E. (2021). Cross-lingual Contextualized Topic Models with Zero-shot Learning. In Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics: Main Volume (pp. 1676–1683). Association for Computational Linguistics.
- https://github.com/MilaNLProc/contextualized-topic-models

Hierarchical stochastic block models

- Gerlach, M., Peixoto, T. P., & Altmann, E. G. (2018). A network approach to topic models. Science Advances, 4(7), eaaq1360. http://doi.org/10.1126/sciadv.aaq1360
- https://github.com/martingerlach/hSBM_Topicmodel

top2vec

- Angelov, D. (2020). Top2Vec: Distributed Representations of Topics. Arxiv. Retrieved from http://arxiv.org/abs/2008.09470v1
- https://github.com/ddangelov/Top2Vec/

BERTopic

- Grootendorst, M., & Reimers, N. (2021). MaartenGr/BERTopic: v0.9.1 (Version v0.9.1). Zenodo. http://doi.org/10.5281/ZENODO.4381785
- https://github.com/MaartenGr/BERTopic

The results of topic modeling are shown in Research.fi

New filter added in the Projects section ("Identified topics"), see image below

New visualizations in the Projects section's 'Show as image' feature (new theme: "Identified topics"), see images below

Top screenshot

Research.fi

Home   Search   Science and Innovation Policy ⌄   Funding calls   Science and research news   EN ⌄

Search target ⌄ | For example, publication, field of science, keyword | 🔍 SEARCH   Search help ⌄

| Publications 584 617 | People Coming soon | Projects 7 403 | Research data 10 436 | Infrastructures 135 | Other research activities Coming soon | Organizations 74 |

Projects - 7 403

Displaying results 1 - 10 / 7403   10   results / page   What project information is included in the service?

[📊 Show as image] 1.

⚙ Filter results

Starting year ⓘ ⌄

＋ Organization ⓘ

Funder ⓘ ⌄

＋ Funding instrument ⓘ

＋ Field of science ⓘ

— Topic ⓘ

Keywords ⓘ ⌄

Themes ⓘ ⌄

Academy of Finland's research fields ⓘ ⌄

Identified topic ⓘ ⌄

PROJECT NAME ⇅	FUNDER ⇅	RECIPIENT ⇅	STARTING YEAR ⇅
URBAN ENVIRONMENT AND CLIMATE CHANGE IN THE ARCTIC: DATA-DRIVEN INTELLIGENCE APPROACH TO MULTIHAZARD MITIGATION	Academy of Finland	Elena Kozlovskaya, University of Oulu	2022
Artificial Intelligence for Twinning the Diversity, Productivity and Spectral Signature of Forests	Academy of Finland	Matti Mõttus, VTT Technical Research Centre of Finland Ltd	2022
Centre of Excellence in Music, Mind, Body and Brain	Academy of Finland	Petri Toiviainen, University of Jyväskylä	2022
Finnish centre of excellence in Randomness and STructures (FiRST)	Academy of Finland	Eero Saksman, University of Helsinki	2022
Uncertainty-aware neural language models	Academy of Finland	Jörg Tiedemann, University of Helsinki	2022
Prediagnostic early synaptic disturbances in neurodegenerative diseases	Academy of Finland	Annakaisa Haapasalo, University of Eastern Finland	2022
Maximally Autonomous AI Assistant	Academy of Finland	Samuel Kaski, Aalto University	2022
Sustainable Drug Discovery and Development with End-of-Life Yield	Academy of Finland	Jari Yli-Kauhaluoma, University of Helsinki	2022

Bottom screenshot

Distribution of projects by year ⓘ Additional information   ☒ Close

[Select theme   ▾] 2.

⚙ Filter results

Starting year ⓘ ⌄

＋ Organization ⓘ

Funder ⓘ ⌄

＋ Funding instrument ⓘ

＋ Field of science ⓘ

— Topic ⓘ

Keywords ⓘ ⌄

Themes ⓘ ⌄

Academy of Finland's research fields ⓘ ⌄

Identified topic ⓘ ⌄

[🗑 Clear filters]

No search word

| Number of projects | Percentage of the number of projects |
| Amount granted | Percentage of the amount granted |

[⬇ Download as image (coming soon)]

(Bar chart: y-axis 0 – 2 000, x-axis years 2013, 2014, 2015, 2016, 2017, 2018, 2019, 2020, 2021, 2022)

Select theme

Funder

Organization

Funding instrument

Field of science

Identified topic **3.**

Funder ⓘ ⌄

\+ Funding instrument ⓘ

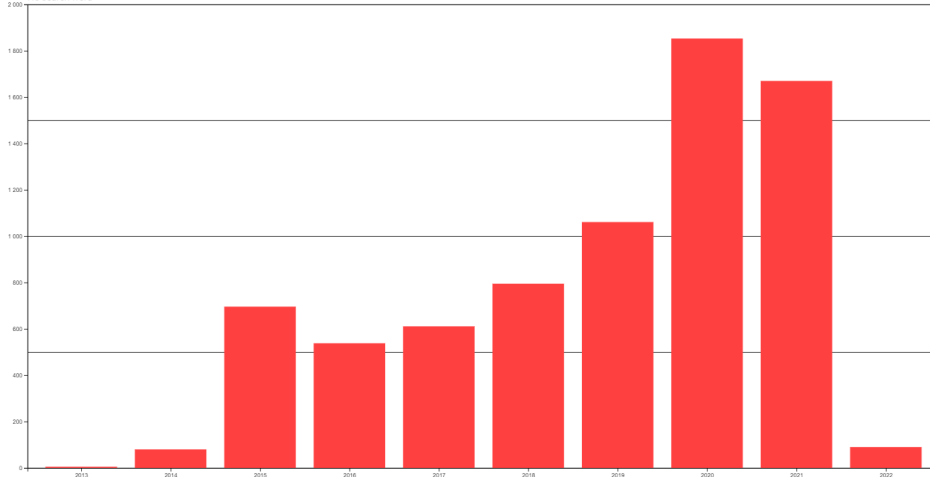\+ Field of science ⓘ

− Topic ⓘ

Keywords ⓘ ⌄

Themes ⓘ ⌄

Academy of Finland's research fields ⓘ ⌄

Identified topic ⓘ ⌄

🗑 Clear filters

No search word

| Number of projects | Percentage of the number of projects |
| Amount granted | Percentage of the amount granted |

⤓ Download as image (coming soon)