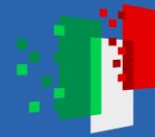




Finanziato  
dall'Unione europea  
NextGenerationEU



Ministero  
dell'Università  
e della Ricerca



Italiadomani  
PIANO NAZIONALE  
DI RIPRESA E RESILIENZA



Consiglio Nazionale  
delle Ricerche

# Text Mining in and for policy making within FOSSR

Antonio Zinilli CNR- IRCrES

Emanuela Varinetti CNR- IRCrES

09 October 2024 – Policy makers session



# FOSSR

Fostering Open Science in Social Science Research  
Innovative tools and services to investigate economic and societal change



Finanziato  
dall'Unione europea  
NextGenerationEU



Ministero  
dell'Università  
e della Ricerca



Italiadomani  
PIANO NAZIONALE  
DI RIPRESA E RESILIENZA



Consiglio Nazionale  
delle Ricerche

## Introduction

This policy brief highlights the **advantages of** using **Text Mining (TM)** to enhance the analysis of **public policies**

TM enables the efficient analysis of large volumes of policy documents, helping to identify trends, issues and impacts of policy decisions.

By integrating innovative tools, data, services and methods, **FOSSR** aims to enhance research practices and policy evaluation in the context of Science, Technology, and Innovation (STI).



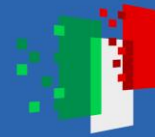
Consortium of European Social Science Data Archives aims to create a centre for excellence in data sharing, enhance the secondary analysis and promote responsible data use in social, economic and environmental studies.



Survey of Health, Ageing and Retirement in Europe, is a research infrastructure for studying the effects of health, social, economic and environmental policies over the life-course of European citizens and beyond.



Research Infrastructure for Science and Innovation policy Studies provides data and services to support the development of analyzes and indicators for the study of science, technology and innovation processes.

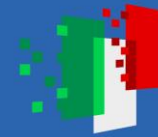


## Text Mining in STI studies

**Text Mining in STI allows responding to specific questions, below is an example:**

- what are the emerging technologies and the most pressing social challenges?
- how are social challenges linked to technological developments?
- how well R&D funding policies align with the Sustainable Development Goals (SDGs)?
- can Text Mining monitor trends over time in environmental and societal themes within research topics?





## The innovative contributions of Text Mining to policy studies

**TM** sheds light on areas overlooked by conventional content analysis and **complement traditional research methods** throughout the policy cycle (Wencker, 2019; Rubinstein et al. 2016)

**Application** of TM in policy contexts and **in policy-making cycle** and emphasis on growing adoption of TM tools, particularly **in the agenda-setting stage**, where are used to support identification and prioritization of policy issues (Ngai and Lee, 2016; Jann and Wegrich, 2007)

Potential of AI tools and machine learning to enhance the entire policy-making process, **improving public sector efficiency** and aiding government decision-making (Berryhill et al. 2019)

Big data analysis accelerates report generation and strengthens **evidence-based policy-making** (Höchtel et al. 2016)

In numerous case studies of various policy areas text mining represents a powerful tool providing deep insights and supporting **data-driven decision-making** (Rubinstein et al. 2016)



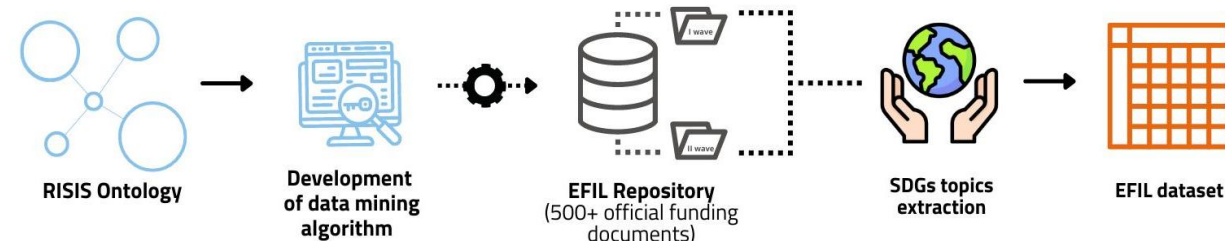
## How FOSSR can address Text Mining in STI context

Potential of this approach in STI policies through two TM applications:

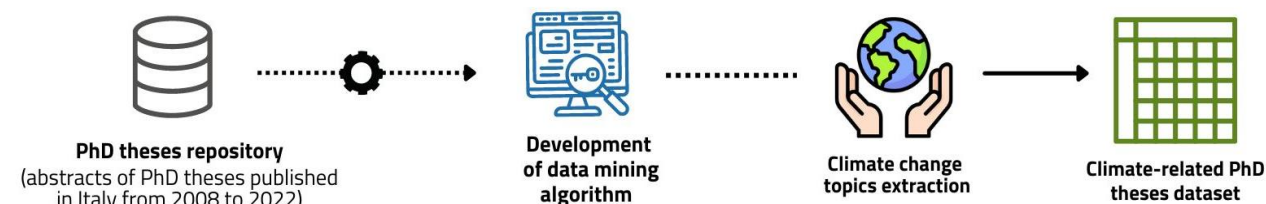
I. Extraction of information about **gender equality** through TM techniques to map the SDG orientation of research funding programs across various Research Funding Organizations (RFOs) in Europe.

II. Use of TM techniques on PhD theses repository published in Italy from 2008 to 2022 to observe orientation and evolution of PhD research topics over time on themes of collective interest, such as **climate change**.

MAPPING GENDER EQUALITY  
ACROSS EUROPEAN RESEARCH  
FUNDING ORGANISATIONS



ANALYSING ORIENTATION  
TOWARDS CLIMATE CHANGE  
IN ITALIAN PHD THESES



\*Icon credits: amonrat, rungrea, fangsoi



## Case 1: SDGs in research funding instruments using EFIL dataset

Can text mining help evaluate the alignment of R&D funding policies with SDGs?

- **EFIL Dataset** (within FOSSR) aims at enabling users to investigate **public R&D funding in Europe** at the level of project funding instruments and Research Funding Organizations (RFO), addressing questions related to policy design and policy implementation (Reale et al., 2023)
- EFIL stores an **official document repository** on RFO funding instruments where an **automated process** (ontologies and statistical analysis) was applied to generate SDGs descriptors (Zinilli et al., 2023)
- TM was applied to over 500 extensive documents to track shifts over time in RFO orientations toward specific SDG goals
- The focus on **gender equality SDG** provides insight into how RFOs align their strategies with sustainability objectives, enabling more precise and impactful policy development

SDG mapping enhances our understanding of **how sustainability targets influence project funding instruments**, aids European countries, as well as RFOs in better integrating SDGs into their research policy agendas.



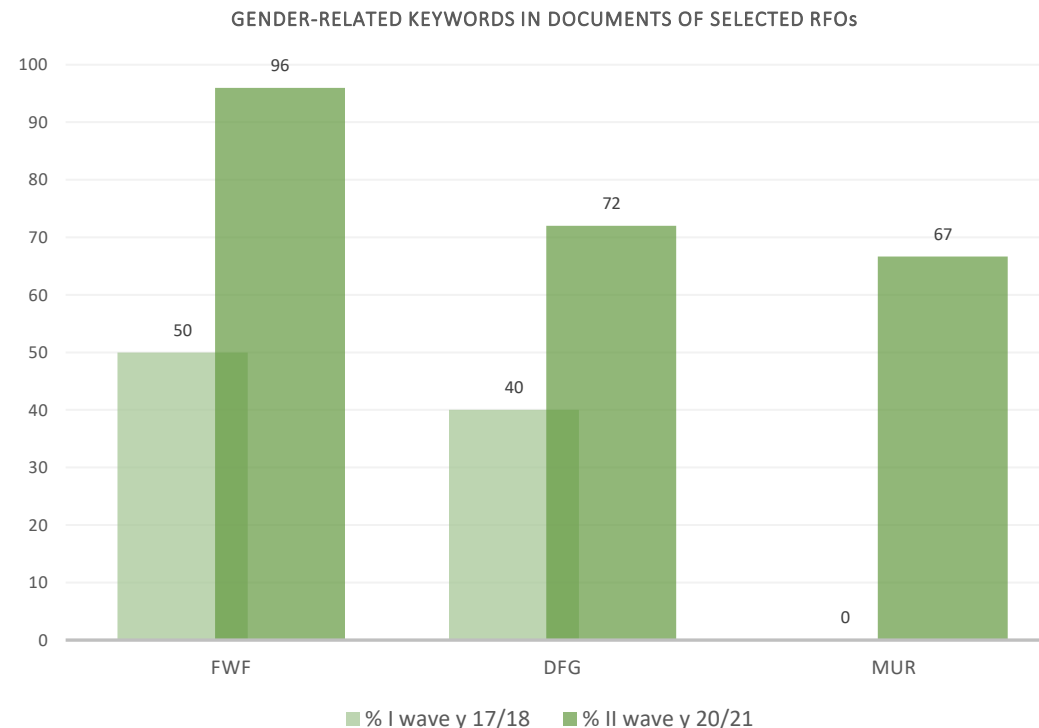


## Case 1: SDGs in research funding instruments using EFIL dataset

### Text Mining to compare integration of gender dimensions across RFOs

Institutions can ensure more equitable opportunities in research funding, as well as **align strategic planning with broader societal goals** using Text Mining to **monitor** and **compare** the integration of gender dimensions across different RFOs and countries.

This **method empowers RFOs to prioritise gender balance in leadership positions, allocation funding and policy development**. Thereby contributing to a richer, more diverse, and inclusive research environment that is responsive to evolving priorities and fosters a more equitable and innovative scientific community.





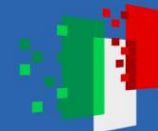
## Case 2: Text Mining applied on Italian PhD thesis

**TM techniques applied on titles and abstracts of PhD theses** dataset in Italy from 2008 to 2022.

Observation of the orientation and evolution of PhD research topics over time on themes of collective interest, such as **climate change**, to understand the **trends in academic research** on critical issues like climate change crucial from a policy perspective:

- Helps policymakers **assess** whether current **research efforts** are aligned with **national and international priorities**.
- enables the **identification of research gaps** that may require targeted funding or policy support to ensure that all important aspects of climate change are being adequately addressed.
- monitors the evolution of research themes, where policymakers can **anticipate emerging challenges** and develop strategies that support **long-term sustainability** and innovation in addressing climate-related **issues**.



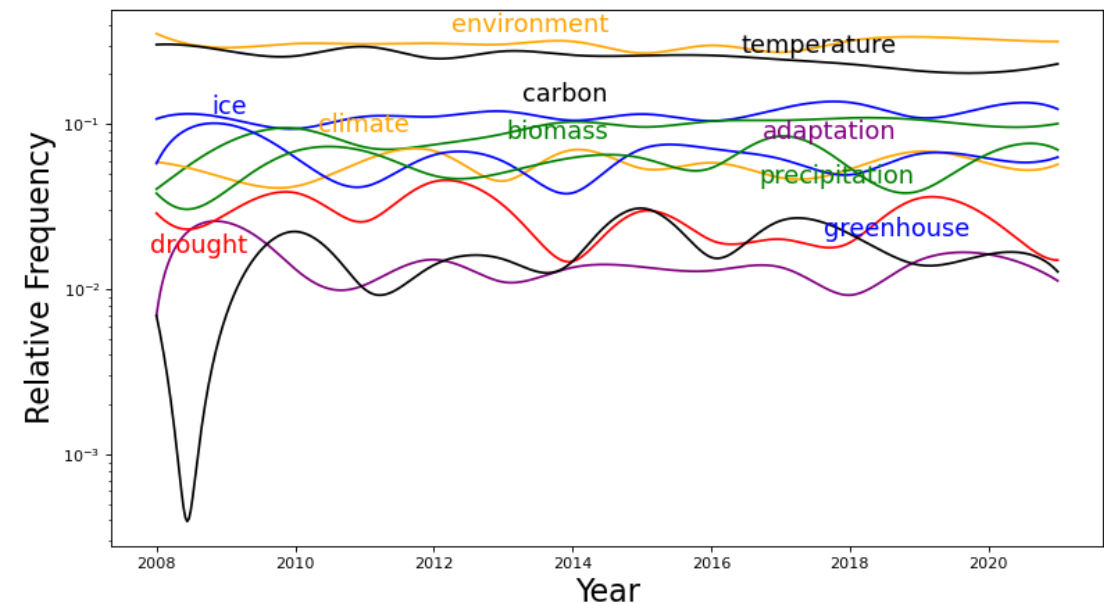


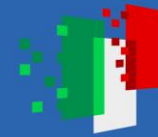
## Case 2: Text Mining applied on Italian PhD thesis

how orientation and emphasis of PhD research on climate-related themes have evolved over time?

Analysis sheds light on how emerging scholars are contributing to the field and whether their research aligns with national and global environmental priorities.

Understanding the trends in PhD research helps policymakers **identify potential gaps in knowledge**, allowing them to craft policies that not only support innovation but also **promote a research setting** more **responsive** to the most urgent **environmental challenges**.





## Case 2: Main Research Topic

### Key areas of PhD research

- Intersections between international cooperation, policy, and environmental impacts
- Research on public health and safety during and after extreme events emphasises the necessity of **integrating climate resilience into health and safety policies**
- Scientific interest of young researchers in understanding the causes and effects of public perception of climate policies
- Developing strategies that can **bridge the gap between scientific evidence and societal attitudes**, ultimately fostering more effective climate action.

This methods could play a pivotal role in cultivating a highly skilled and knowledgeable workforce, equipped to lead advancing climate policy and sustainability initiatives.

#### Main Research Topic

Climate Change and Human Impact

International Cooperation and Climate

Climate Change and Policy

Pollution and Ecosystems

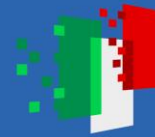
Public Health and Safety in Crises

Industry, Pollution and Environmental Impact

Global Warming

Climate Change: hoax or reality

Meat Consumption, Agriculture, and Environmental Impact



## Policy evidence

**Stakeholders** are encouraged to consider Text Mining as a powerful tool for enhancing evaluation processes across various sectors. By employing Text Mining techniques, **policy makers** can gain deeper insights, enabling more informed decision-making and improved outcomes.

Through **FOSSR services**, policy makers can base their strategies on concrete, timely evidence, enhancing their ability to respond to complex and continuously evolving challenges like climate change or gender inequality.

It enables the **exploration of interconnected policy areas**, revealing how issues like **climate change and social justice** are addressed in an integrated manner across research and global debates.

TM can identify subtle changes in the language used in policy documents or scientific outputs, signaling a potential shift in priorities.

TM allows for ongoing **monitoring of policy effectiveness**, promoting a more dynamic and **adaptive decision-making process** compared to traditional approaches.

TM not only increases transparency in policy-making but also helps policymakers **detect weak signals**—emerging trends or concerns that may **not be immediately visible through other methods**.



## References

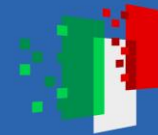
- Albert, T., Moehrle, M. G., & Meyer, S. (2015). Technology maturity assessment based on blog analysis. *Technological Forecasting and Social Change*, 92, 196-209.
- European Commission: Competence Centre on Text Mining and Analysis. [https://knowledge4policy.ec.europa.eu/text-mining\\_en](https://knowledge4policy.ec.europa.eu/text-mining_en) (2016).
- Höchtl, J., Parycek, P., & Schöllhammer, R. (2016). Big data in the policy cycle: Policy decision making in the digital era. *Journal of Organizational Computing and Electronic Commerce*, 26(1-2), 147-169.
- Jann, W., & Wegrich, K. (2007). Theories of the policy cycle. In F. Fischer, G. Miller, & M. S. Sidney (Eds.), *Handbook of public policy analysis: Theory, politics, and methods*. CRC Press. <https://doi.org/10.4324/9781315093192>
- Ngai, E. W., & Lee, P. T. Y. (2016). A review of the literature on applications of text mining in policy making.
- Reale, E., Spinello, A. O., Varinetti, E., & Zinilli, A. (2023). Documentation of RISIS datasets: EFIL (2.0). Zenodo. <https://doi.org/10.5281/zenodo.7938134> (Citing a website)
- Rubinstein, M., Meyer, E., Schroeder, R., Poel, M., Treperman, J., van Barneveld, J., ... & Svetachova, M. (2016). Ten use cases of innovative data-driven approaches for policymaking at EU level. Report prepared in the context of: Data for Policy: A study of big data and other innovative data-driven approaches for evidence-informed policymaking.
- Wencker, T. (2019), Text Mining in Evaluation, DEval Policy Brief 1/2020, Deutsches Evaluierungsinstitut der Entwicklungszusammenarbeit (DEval), Bonn.
- Zinilli A. Reale E., Spinello A.O. Varinetti E. (2023). Diversification vs specialization from the perspective of research programmes: A complexity approach. Atlanta Conference On Science and Innovation Policy. MAY 24-26, 2023. ATLANTA, GEORGIA, USA GEORGIA TECH GLOBAL LEARNING CENTER



Finanziato  
dall'Unione europea  
NextGenerationEU



Ministero  
dell'Università  
e della Ricerca



Italiadomani  
PIANO NAZIONALE  
DI RIPRESA E RESILIENZA



Consiglio Nazionale  
delle Ricerche

# THANK YOU!



[fossr.dissemination@ircres.cnr.it](mailto:fossr.dissemination@ircres.cnr.it)



[@fossrproject](https://twitter.com/fossrproject)



[fossr.eu](https://www.facebook.com/fossr.eu)



[fossr-eu](https://www.linkedin.com/company/fossr-eu)



[@fossr](https://www.youtube.com/@fossr)



[zenodo.org/communities/fossr](https://zenodo.org/communities/fossr)