T13P04 / Advanced Computational Methods for Public Policy: NLP, ML, and LLMs

Topic : T13 / METHODOLOGIES

Chair : Antoine Lemor (Université de Sherbrook (QC, Canada) ; Réseau francophone international en conseil scientifique (RFICS) ; Centre de recherche sur la science et la technologie (CIRST))

Second Chair : Louis-Robert Beaulieu-Guay (University of Saskatchewan)

Third Chair : Igor Tkalec (University College London, University of London)

GENERAL OBJECTIVES, RESEARCH QUESTIONS AND SCIENTIFIC RELEVANCE

Context

Natural Language Processing (NLP) is a subfield of computer science and linguistics that focuses on enabling computers to understand and interact with human language. It uses techniques from Machine Learning (ML), a branch of artificial intelligence that allows computers to learn from data and improve their performance. In the context of policy analysis, ML-driven NLP can be used for a variety of tasks, such as automatically categorizing policy documents or identifying policy change. Recently, the development of Large Language Models (LLMs), such as GPT-4, has further advanced NLP capabilities. LLMs are models trained on vast amounts of text data, allowing them to comprehend and generate human-like language. LLMs can assist in many tasks both in policy analysis and policy advice through their ability to process large quantities of data.

Natural Language Processing (NLP) and Machine Learning (ML) are becoming a powerful tool in policy analysis. With the increasing digitization of policy documents, these methods have been applied to various types of texts—ranging from legislative documents and regulations to policy briefs and political debates—enhancing the efficiency and scope of policy analysis. However, while NLP and ML show potential in automating routine tasks, debates continue about whether it can fully replace human policy analysts, especially when it comes to the interpretation of context or complex policy document and ethical considerations (Safaei & Longo, 2024; Bol & Bono, 2024). **How can NLP, enriched by ML and LLMs, enhance public policy analysis, and what are their limitations? This panel aims to critically examine the (1) types of data used, (2) the methods employed—including the roles of ML and LLMs—and (3) the applications of these technologies in public policy, exploring both their current use and future potential.**

(1) Types of textual data

The effectiveness of NLP and ML in policy analysis depends on the availability and quality of textual data. Current efforts include large corpora of parliamentary proceedings, such as the ParlaMint dataset, which contains transcriptions of European parliamentary debates (Erjavec et al., 2023). Similarly, RegCan, a database of Canadian federal regulations, provides data for studying regulations (Tremblay-Faulkner et al., 2018). Policy briefs, party manifestos, and political speeches are also being analyzed using NLP techniques (Orellana & Bisgin, 2023). These textual datasets provide the raw material for NLP and LLMs, but they also present challenges due to the complexity of legal and political language, preprocessing, or inherent bias, requiring advanced methods and controls to use and analyze these documents accurately.

(2) Types of methods used

Various NLP methods are applied to analyze these textual data. Text classification techniques categorize documents into policy areas such as health, education, or climate (Jin & Mihalcea, 2023). Topic modeling is another common method, which identifies latent themes within large corpora, analyzing trends in policy debates (Singh et al., 2024). Sentiment analysis measures the emotional tone in parliamentary speeches (Rheault et al., 2016). More advanced methods, such as causal mapping and graph analytics, allow for the extraction of cause-effect relationships from policy documents, enabling the visualization of policy impacts and interdependencies (Hooper et al., 2024). Machine Learning algorithms can enhance these methods by enabling predictive analysis and pattern recognition. LLMs also can perform tasks like summarization, question-answering, and even drafting policy documents. However, despite these advancements, concerns about model interpretability, ethical use, and biases persist (Rodriguez et al., 2024).

(3) Types of applications in policy analysis and policymaking

The applications of NLP for policy are broad, ranging from policy evaluation to the production of policy briefs. Automated systems and LLMs now generate policy briefs by summarizing legislative texts (Safaei & Longo, 2024). NLP tools are also used to monitor and evaluate the implementation of international climate agreements (Singh et al., 2024). Furthermore, sentiment analysis of public opinion data—such as social media—enables policymakers to assess public reactions to policy changes (Rodriguez et al., 2024). Despite these advances, the idea that LLMs could entirely replace human policy analysts remains controversial (Bol & Bono, 2024).

CALL FOR PAPERS

We invite contributions focusing on: (1) the types of textual data analyzed, (2) the methodologies employed—including NLP, ML, and LLMs—and (3) the applications in policy analysis/policymaking. While these technologies are transforming how researchers and policymakers analyze large bodies of data, challenges remain, particularly in interpreting complex policy documents and ensuring ethical use. This panel seeks to address these issues, bringing together research on the application of NLP, ML, and LLMs in public policy and exploring their potential for future development.

We are particularly interested in papers that address one or more of the following themes:

(1) Types of textual data used for policy analysis

Submissions may explore the variety of textual data being used in current NLP, ML and LLMs applications for policy analysis. Relevant data sources include legislative texts, regulations, policy briefs, political manifestos, speeches, and parliamentary debates. Papers can discuss how these datasets are processed, the challenges of working with complex legal or political language, and the potential for expanding the range of data used in policy analysis.

(2) Methods applied

We welcome papers that present or evaluate different methodologies for analyzing policy-related data. Methods of interest include text classification, topic modeling, sentiment analysis, causal mapping, predictive analytics, and the use of LLMs for advanced language tasks. Contributions may focus on how these techniques are used to identify policy trends, predict policy outcomes, map relationships between policy issues, or assess the sentiment and tone of political discourse. Papers can also address limitations and ethical concerns, such as biases within models, transparency, interpretability, and the challenges of deploying ML and LLMs in policy contexts.

(3) Applications in policymaking

Submissions may highlight the practical applications of these technologies in various aspects of policymaking. These could include the automated production of policy briefs using LLMs to summarize legislative texts, real-time monitoring and evaluation of policy implementation, or predictive modeling to forecast policy outcomes. Papers are encouraged to explore how these applications can support decision-making processes, improve policy design, or enhance the understanding of complex policy issues. We are also interested in critical assessments of whether these technologies can replace or should only complement human analysts and policymakers.

We encourage contributions from both researchers and practitioners who are exploring the intersection of NLP, ML, LLMs, and public policy. Submissions may present empirical studies, methodological innovations, case studies, or theoretical reflections. The goal of this panel is to foster an interdisciplinary dialogue on how these technologies are reshaping the field of public policy, the practical challenges they face, and the opportunities for further development.

T13P04 / Advanced Computational Methods for Public Policy: NLP, ML, and LLMs

Chair : Antoine Lemor (Université de Sherbrook (QC, Canada) ; Réseau francophone international en conseil scientifique (RFICS) ; Centre de recherche sur la science et la technologie (CIRST))

Second Chair : Louis-Robert Beaulieu-Guay (University of Saskatchewan)

Third Chair : Igor Tkalec (University College London, University of London)

Session 1Advanced computational methods for political science and public policy

Wednesday, July 2nd 13:45 to 15:45 (B9)

Discussants

Antoine Lemor (Université de Sherbrook (QC, Canada) ; Réseau francophone international en conseil scientifique (RFICS) ; Centre de recherche sur la science et la technologie (CIRST))

Louis-Robert Beaulieu-Guay (University of Saskatchewan)

Applied machine learning for political science: an overview

Igor Tkalec (University College London, University of London)

On the premise that machine learning (ML) might advance quantitative political science, this article reviews to what extent and for what purpose political science analyses utilise ML modelling techniques. The article analyses academic articles published in selected courpus of political science journals that applied ML modelling. The initial expectation is that applied ML is on the rise in political science, especially in regard to quantitative text aanlyses. In addition, it addresses a mixed- or multi-method potential to be realised through combining ML modelling with other quantitative and qualitative techniques convetionally across political science. Moreover, the analysis examines practices on how ML models applied to political science questions are presented and interepreted.

Generative Artificial Intelligence as a Canadian Policy Expert: Results from Advanced Fine-tuning and Customization of a Base Large Language Model

Stany Nzobonimpa (École nationale d'administration publique, Université du Québec)

Jean-Francois Savard (ENAP)

Justin Lawarée (École nationale d'administration publique, Université du Québec)

Recent literature has demonstrated that while Generative Artificial Intelligence may not be the cream of the crop for policy analysis exercises, its potential should not be undermined and various models have demonstrated notable competencies for policy-related work including in the field of crafting policy briefing notes, writing legislation, and decoding complex policy materials. Researchers have consistently pointed out this advantage and suggested specialized models may be the way to go for model improvement. In this paper, we pursue the same challenge. By following up on an earlier paper we published in AI & Society in late 2024, we seek to leverage advanced techniques for model fine-tuning to customize a base Large Language Model for policy specialization in the context of the Canadian federal public sector. We answer the following question: Does fine-tuning a base model with past policy data make it better than a general-purpose foundation model? To answer the question, we designed an experimental study in which we leverage LLM fine-tuning capabilities of the Python programming language through Google’s Colab project and create a policy-specialized model that we deploy in the form of an experimental chatbot. Specifically, we use fine-tuning techniques offered to Python developers through the OpenAI API service and packages to create a policy-trained model that is able to respond to users’ inquiries on policy-related subjects. The chatbot is then put to test with a sample of human users made of policy analysts, advisors and executives who rate the results and policy crafting capabilities of the fine-tuned model using a framework

and an index published in the earlier work.

(Virtual) Party manifestos and parliamentary questions: What do they say about environment and environmental politics and policies in India?

Pradeep Guin (Jindal School of Government and Public Policy)

Papiya Mazumdar (University of Leeds)

Background: Our environment is slowly and continuously degrading. Safeguarding our environment form anthropogenic onslaughts, therefore, should be of paramount importance. This is more so in light of increasing number of devastating hazards that are linked to climate change, which must be a significant source of worry. India's development narrative has most recently been favouring infrastructure development-led economy over protecting our environment. Poor global rankings – such as 176 out of 180 in the 2024 Environmental Performance Index (EPI), and till most recently being the seventh-most vulnerable country globally to climate change, and number one in terms of fatalities and losses – leaves us with no choice but ask, where as a country are we heading to? Will we leave the kind of natural environment that we have today, for our future generations? In this study, we aim to explore the status of environment and environmental policymaking in India through the lens of election manifestos and parliamentary debates.

Data and methods: We use digitised text data from two primary sources: (a) election manifesto documents of various political parties across all general elections (1951–2024), and (b) parliamentary debate data (1999–2019). While we have access to all the manifestos of three major political parties, representing three ideologies (left, centre, and right), we are in the process of obtaining manifestos of remaining prominent parties whose members represent the Indian Parliament (20 parties). We obtained parliamentary data from the questions that were asked by the elected parliamentarians in the Lok Sabha (lower house of India's bicameral Parliament).

Our analysis is a work-in-progress. We have applied, and will continue to explore, machine learning (ML) and natural language processing (NLP) tools to analyse our data, including performing sentiment analysis on both data types. For now, we have used Google Colab to code using several libraries within Python (pandas, numpy, matplotlib.pyplot) to produce frequency distribution and graphs of environment-related words. Apart from the artificial intelligence (AI)-based analysis we perform other statistical analysis to explore developments in the environment policymaking space in India. This includes studying various characteristics of the questions raised in the parliament that are related to environment and environmental issues. This includes political party representation of the parliamentarian, their gender, constituency and its type (SC, ST, general), and whether members are from the treasury or opposition bench.

Preliminary results and contribution to literature: Frequency distribution of environment-related words is a strong proxy indicator of the degree of seriousness demonstrated by political parties towards environment. Our analysis would help systematically identity the discourses that have taken place on environmental agenda over time in India using AI-based analysis. An objective understanding on this matter will be foundational to future pro-environmental agenda setting and policy implementation in India. This would be essential for us to leave a healthy place for our future generations.

T13P04 / Advanced Computational Methods for Public Policy: NLP, ML, and LLMs

Chair : Antoine Lemor (Université de Sherbrook (QC, Canada) ; Réseau francophone international en conseil scientifique (RFICS) ; Centre de recherche sur la science et la technologie (CIRST))

Second Chair : Louis-Robert Beaulieu-Guay (University of Saskatchewan)

Third Chair : Igor Tkalec (University College London, University of London)

Session 2NLP applications in policy and public administration contexts

Thursday, July 3rd 08:00 to 10:00 (D7)

Discussants

Igor Tkalec (University College London, University of London)

Evidence and Influence: A Systematic Analysis of Consultation Practices in Regulation

Louis-Robert Beaulieu-Guay (University of Saskatchewan)

Antoine Lemor (Université de Sherbrook (QC, Canada) ; Réseau francophone international en conseil scientifique (RFICS) ; Centre de recherche sur la science et la technologie (CIRST))

This project examines the impact of public consultations on the Canadian regulatory process, a key but understudied area
of Canadian public administration. Public consultations, often referred to as “notice-and-comment” processes in regulatory
contexts, are formal mechanisms through which government agencies invite feedback from individuals, organizations, and
other stakeholders on proposed regulations or policies. Although these consultations are crucial opportunities for civil
society to influence policy, their actual impact on final decisions remains controversial. Some studies suggest that they
amplify the influence of already dominant economic actors, while others indicate that they can inform regulators and
enhance the quality of policies. This research aims to clarify these issues by providing a systematic and empirical analysis
of the concrete impact of regulatory consultations in Canada since 2012. It identifies the types of actors, stakeholders and
arguments that do influence regulators, while assessing whether public consultations enable better integration of external
knowledge into policymaking. Involving a broad range of voices can not only improve policy outcomes but also help identify
potential issues or unintended consequences, and ultimately increase trust in the policy-making process. This project
directly addresses these goals.

This research will make a contribution to the literature on regulatory governance and participatory processes in
two main ways. (1) First, by leveraging recent and rigorous methodological approaches that combine advanced Natural
language processing (NLP) techniques, statistics, and manual coding, the project will systematically identify and analyze—
drawing on a large textual database—the changes between regulatory proposals before and after public consultations. In
doing so, it will reveal the concrete influence of all relevant stakeholders and the arguments they put forward. It will
represent the first comprehensive study to quantify the impact of “notice-and-comment” public consultations

on the
Canadian regulatory corpus in a reproducible manner, using an analytical framework that can be applied to other studies
across different contexts and time periods. (2) Most importantly, the findings will deepen our understanding of the
relationships between regulators and stakeholders by identifying the conditions under which consultations can—and do—
influence regulation. These results can also inform broader debates about the role certain actors play in shaping public
policies, thereby helping to improve Canadian policies.

Credible Science, Influential Science? A Computational Study on the Influence and Credibility of Public Scientific Research Agencies in Public Health in Canada

Antoine Lemor (Université de Sherbrook (QC, Canada) ; Réseau francophone international en conseil scientifique (RFICS) ; Centre de recherche sur la science et la technologie (CIRST))

This article aims to examine the influence of public scientific research agencies (PSRAs) on public policy in Canada and Quebec, using the Public Health Agency of Canada (PHAC), the National Institute of Public Health of Quebec (INSPQ), and the World Health Organization (WHO) as case studies. In a context where public issues such as climate change and epidemics are particularly complex, these agencies play a key role by producing scientific evidence that can inform policymaking and the public. This study employs natural language processing (NLP) methods designed to model the influence of these agencies using multiple textual databases. The study seeks to measure the influence of PSRAs by analyzing three types of relationships: (1) the indirect influence of PSRAs through public debate via the media; (2) their direct influence in parliamentary debates; and (3) their tangible influence, demonstrated by the adoption of legislative texts. The main hypothesis is that the influence of PSRAs on public policy is linked to their credibility, which depends on three key factors: the affective and political context (including uncertainty and polarization), the type of evidence produced (its robustness and disciplinary origin), and the domain expertise relevant to the issue at hand. If PSRAs enjoy high credibility among the media and policymakers, they are more likely to influence policy. To test this hypothesis, the study employs a mixed-methods approach, using machine learning techniques to analyze two textual databases: one comprising articles published by 1,905 media organizations since 2017, and the other containing transcripts of federal and provincial legislative processes. This study aims to develop a unified theory of the factors of influence of experts and evidence on public policy in democratic contexts, while integrating the essential role of the media as intermediaries between science and politics. It thus contributes to understanding the science-policy interface and improving decision-making by providing a reproducible analytical framework to assess the influence of PSRAs on public policy.

Protestor Partisanship and Police Response: Evidence from France, Germany, and the United Kingdom

Katelyn Nutley (University of St Andrews)

Anecdotal evidence from a number of liberal democracies suggests certain groups, largely ‘Leftist’, are unable to exercise their democratic rights for fear of reprisal. Due to recent reports, this article empirically investigates how protestor partisanship impacts police response in France, Germany, and the United Kingdom. I use the Armed Conflict Location and Event Dataset’s (ACLED) observational protest data, which identifies political actors and has detailed event descriptions, as a jumping-off point. Around this list of political actors, a novel, multinational dataset on the policy positions of various direct actions groups is built. Through the use of a supervised machine learning technique, these preferences are used to assign protestors classifications of: “left.” “centre left”, “unknown”, “centre right”, and right (Grimmer, Roberts, and Stewart, 2021). Given the obvious dimensionality of political space, a variable that categorises the extent to which a group is aligned with a party or otherwise represented in government is also included. By relying on text classification methods, the protest data is then automatically supplemented with a binary and categorical measure denoting police response and severity, respectively (insert citation here). As the observations of police behaviour in ACLED’s data do not reach the requisite level of statistical power, this chapter of the dissertation leans on non-parametric Bayesian estimation methods (also called imputation) to fill the gaps created by limited or missing data (insert citation). Through the use of logistic regression models, this article models the country-specific likelihood that the Left, loosely and strictly defined, will be over-policed and that the Right will be under-policed during protest events.

Mission-Driven Performance: An NLP Approach to Performance Dynamics

Hyemin Lee (Ewha Womans University, Republic of Korea)

Yujin Choi (Ewha Womans University)

The performance of public institutions and its determinants has long been recognized as a critical issue in the field of public administration. However, existing studies predominantly emphasize resource-oriented aspects, such as staffing changes and financial management, without adequately addressing the role of missions. Mission statements are essential for public institutions, providing purpose, guiding decisions, and aligning goals. Shaped through interactions between stakeholders and members, they influence resource allocation, motivate performance, and build trust by connecting work to societal values (Rainey & Steinbauer, 1999; Alegre et al., 2018).

Rainey and Steinbauer's (1999) theoretical framework highlights that organizational performance emerges from the interplay of diverse factors, including resources, stakeholders, organizational culture, leadership, and mission. Despite this nuanced perspective, mission-focused research remains limited, particularly in the Korean context. Each institution's unique mission defines its purpose and objectives, suggesting that overlooking its significance can fundamentally compromise organizational success.

This study seeks to bridge this research gap by empirically examining how mission statements of public institutions in South Korea impact organizational performance. Moreover, we investigate the mission's formal role and its substantive influence through organizational members' understanding, thereby providing a comprehensive analysis of mission dynamics.

Utilizing mission statements directly collected from approximately 200 Korean public institutions, data from the Korean Public Service Organization Survey (KPSOS), and performance evaluation data assessed by the Korean Ministry of Economy and Finance (MOEF), our study applies NLP techniques to analyze missions, inspired by Baek et al. (2023) and Brysbaert et al. (2014). Text analysis identifies key attributes of the mission statements, which are then linked to organizational performance metrics through regression analysis. The findings demonstrate how mission attributes enhance performance and influence employees' perceptions of purpose and motivation. The study concludes by discussing theoretical and practical implications, highlighting the multifaceted role of missions in shaping public institution performance.

References

Alegre, I., Berbegal-Mirabent, J., Guerrero, A., & Mas-Machuca, M. (2018). The real mission of the mission statement: A systematic review of the literature. Journal of Management & Organization, 24(4), 456-473.

Baek, Y. M., Ihm, J., & Kang, C. H. (2023). Does mission concreteness make a difference in nonprofit performance? Automated text analysis approach to the importance of concrete mission statements. Nonprofit Management and Leadership, 34(2), 409-431.

Brysbaert, M., Warriner, A. B., & Kuperman, V. (2014). Concreteness ratings for 40 thousand generally known English word lemmas. Behavior Research Methods, 46(3), 904–911.

Rainey, H. G., & Steinbauer, P. (1999). Galloping elephants: Developing elements of a theory of effective government organizations. Journal of public administration research and theory, 9(1), 1-32.