Translating Natural Language Processing Frameworks in Humanistic Data Analysis

Shi Xian Chng[*1] **Kenneth Y T Lim**[*1] 0000-0003-3756-6625
[1]*Nanyang Technological University, Singapore*

Correspondence to: Kenneth Y T Lim kenneth.lim@nie.edu.sg

1. Introduction

This paper presents a technical investigation into the extent to which natural language processing (NLP) frameworks developed for political analysis can inform data analysis and automation in the Humanities. Focusing on Singapore as a contextual anchor while drawing on generalizable principles, we evaluate the cross-domain applicability of two widely used NLP methods: lexicon-based sentiment analysis and topic modelling via Latent Dirichlet Allocation (LDA).

We apply Liu's opinion lexicon to assign sentiment polarity and LDA to extract latent thematic structures across two contrasting corpora. The first is a political corpus consisting of parliamentary speeches from the ParlLawSpeech dataset [1], and the second is a Humanities corpus composed of selected literary and philosophical texts from Project Gutenberg [2]. By comparing model behavior across these domains, we examine differences in performance, interpretability, and representational assumptions.

Our analysis demonstrates that NLP frameworks designed for political discourse implicitly rely on semantic stability, rhetorical polarity, and relatively discrete thematic boundaries. These assumptions align well with political texts, which tend to employ explicit evaluative language and goal-oriented framing. In contrast, Humanities texts are characterized by metaphorical language, polysemy, interpretive openness, and layered meaning, leading to systematic degradation in sentiment and topic modelling performance. Sentiment polarity becomes muted or misleading, while topic models produce lower coherence and reduced conceptual clarity.

Building on this comparative evaluation, we present a technical failure analysis that identifies key representational mismatches between standard NLP models and Humanities discourse. We argue that these mismatches have important implications for the design of NLP-driven Artificial Intelligence in Education (AIED) systems, particularly those intended for Humanities learning materials. The paper concludes by outlining a design framework that highlights the need for domain-sensitive representations, interpretability-aware modelling choices, and pedagogically aligned evaluation criteria when adapting NLP tools to Humanities contexts.

2. Methodology

This study evaluates the cross-domain transferability of NLP frameworks developed for political analysis when applied to Humanities texts. We focus on two widely adopted methods - Latent Dirichlet Allocation (LDA)(with the equation as seen in Equation (1) for topic modelling and lexicon-based sentiment analysis (Equation (2))

- and compare their performance across two contrasting corpora: parliamentary speeches and literary texts. Our main contribution is a technical failure analysis that identifies representational mismatches underlying performance degradation in Humanities applications, with implications for NLP-driven AIED system design.

2.1 Corpora and Experimental Design

We constructed two domain-specific corpora. First, a Political Corpus from parliamentary speeches from the ParlLawSpeech dataset [1], comprising millions of full-text speeches from European legislatures. This corpus reflects institutional political discourse with stable terminology, explicit rhetorical framing, and well-defined policy themes.

Second, a Humanities Corpus from a curated selection of literary and philosophical works from Project Gutenberg [2], spanning fiction, poetry, and essays. These texts exhibit metaphorical language, polysemy, and interpretive openness typical of Humanities discourse.

To ensure comparability, texts were normalized for length where possible, and analyses controlled for corpus size effects [3]. Evaluation combined quantitative metrics (sentiment distributions, topic coherence) with qualitative inspection of word lists and topic overlap.

2.2 Equations

$$P(\boldsymbol{\theta}, \boldsymbol{z}, \boldsymbol{w} \mid \alpha, \beta) = \prod_{d=1}^{M} P(\theta_d \mid \alpha) \left(\prod_{n=1}^{N_d} P(z_{d,n} \mid \theta_d) P(w_{d,n} \mid z_{d,n}, \beta) \right)$$

Equation (1): Joint probability distribution and genera- tive process of LDA

$$S(d) = \frac{1}{N_d} \sum_{i=1}^{N_d} p(w_i) \, \mathbb{I}(w_i \in L)$$

Equation (2): Sentiment score equation

3. Discussion

Political discourse is inherently strategic and goal-oriented. Politicians employ language to persuade, mobilize, or legitimize positions, often framing issues in explicitly positive or negative terms. As a result, political texts tend to use a relatively constrained and domain-specific vocabulary centered on policy, governance, and power. This rhetorical style "demands strong polarity": speeches advocating economic reform emphasize positively valenced terms such as growth, success, or prosperity, while oppositional rhetoric foregrounds negative terms such as crisis, failure, or threat. Lexicon-based sentiment analysis is well suited to capturing these overt contrasts, and our results align with prior work [1] showing that sentiment signals in political texts often correspond meaningfully to stance, ideology, or framing.

In contrast, Humanities texts prioritize nuance, ambiguity, and aesthetic expression. Literary and philosophical works frequently convey sentiment implicitly through metaphor, irony, narrative voice, or complex syntactic structures. Words flagged as positive or negative by sentiment lexicons may be embedded in contexts that invert or complicate their apparent polarity. For instance, negatively valenced terms may be used approvingly in aesthetic or critical contexts, while positive terms may function ironically. These characteristics pose well-documented challenges [3] for automated sentiment analysis, particularly lexicon-based approaches that rely on surface word counts without contextual interpretation. Our findings reflect this limitation: sentiment analysis struggled to assign stable or interpretable polarity scores to Humanities texts, especially where negation, ambivalence, or layered rhetoric was prominent.

Taken togther, these findings suggest both promise and caution in transferring NLP methods developed in political science to Humanities contexts. On the one hand, even relatively simple techniques such as sentiment lexicons and LDA can surface broad patterns in Humanities corpora. They can help identify recurring concepts, shifts in tone across a text, or thematic emphases across collections of works. In educational settings, such tools may support exploratory analysis, enable large-scale comparisons, or provide entry points for students engaging with complex texts.

4. Conclusion

This study demonstrates that NLP frameworks developed for political analysis do not transfer straightforwardly to Humanities texts. While lexicon-based sentiment analysis and LDA perform robustly on political discourse, their effectiveness degrades systematically when applied to literary and philosophical materials. The observed differences are not incidental but stem from fundamental contrasts in language function: political discourse prioritizes persuasion, evaluation, and policy clarity, whereas Humanities texts emphasize ambiguity, interpretation, and symbolic meaning.

Looking forward, effective NLP support for Humanities learning will require domain-aware adaptations, including context-sensitive representations, word sense disambiguation, and models that capture hierarchical or overlapping thematic structures. Rather than replacing interpretive practices, NLP tools should be designed to augment critical reading and reflection.

Acknowledgments

We would like to acknowledge the funding support from Nanyang Technological University – URECA Undergraduate Research Programme for this research project.

References

[1] Rauh, C., Schwalbach, J.: The ParlLawSpeech Dataset: Full-text parliamentary speeches from 1947 to 2020. Legislative Studies Quarterly 45(3), 1–15 (2020). https://doi.org/10.1111/lsq.12279

[2] Project Gutenberg, "Project Gutenberg Digital Library," https://www.gutenberg.org, accessed Jan. 2026.

[3] A. Schofield and D. Mimno, "Comparing Apples to Apple: The Effects of Stemmers on Topic Models," Transactions of the Association for Computational Linguistics, vol. 4, pp. 287–300, 2016. https://aclanthology.org/Q16-1021/date.