

Foundations for Territorial Disambiguation in Law: A Preliminary Study Using the Education Act 2005

Safia KANWAL ^{a,1}, Livio ROBALDO ^b and Hafsa DAR ^b and Davide LIGA ^c and Joseph ANIM ^a

^aSwansea University, UK

^bUniversity of Gujrat, PK

^cUniversity of Luxembourg, Luxembourg

Abstract. In the devolved legal system of the United Kingdom (UK), legislative provisions may apply differently across regions such as England, Wales, Scotland, and Northern Ireland. Accurately determining this territorial scope is essential for legal interpretation and AI-assisted legal tools. However, metadata capturing jurisdictional applicability is inconsistently format, as only a few Acts include Territorial Application Annexes. This study presents a case study using the Education Act 2005 to evaluate the accuracy of automated methods for identifying territorial scope. We found that only 46.9% of sections matched in jurisdictional coverage. The best-performing approach achieved nearly 80% accuracy, showing that LLMs can effectively support scalable and explainable territorial disambiguation.

Keywords. Territorial Disambiguation, Legal—DocML, Large Language Models

1. Introduction

Legislation in the UK operates under a devolved structure, dividing authority between England, Wales, Scotland, and Northern Ireland. Determining where a provision applies—its territorial scope—is vital for legal interpretation and compliance automation. The National Archives provides statutes in *Akoma Ntoso XML (LegalDocML)* format, using `<extent>` tags to indicate territorial coverage. However, practical application may be narrower: a section marked as extending to England and Wales (E+W) may, in effect, apply only to England, depending on institutional references or legal context. This leads to ambiguity when `<extent>` alone is used for territorial inference. The *Education Act 2005*, which includes a detailed Territorial Application Annex, provides a unique ground truth to evaluate and improve automated territorial scope extraction using AI-based methods. As highlighted in the paper, many Acts lack such annexes, making territorial classification a manual and non-scalable task. This motivated the exploration of

¹Corresponding Author: Livio Robaldo, livio.robaldo@swansea.ac.uk.

AI-driven inference techniques that can mirror human-style reasoning by analysing legal structure, content cues, and institutional references. [1]

2. Methodology

This study used the Education Act 2005 as a benchmark to evaluate automated approaches for territorial classification. Each section of the Act was extracted from its LegalDocML representation and paired with its corresponding ground-truth jurisdiction(s) from the Territorial Application Annex. The dataset was structured to include section text, hierarchical context (Part/Chapter titles), and `text` metadata. Four progressively enriched methods were evaluated: (1) Extent-tag baseline, which directly mapped `text` territories to predicted scope; (2) Named Entity Recognition (NER), which used jurisdictional mentions to refine classification; (3) LLM-based inference, where large language models (DeepSeek-R1 8B and 32B) were prompted using structured context to infer territorial scope; and (4) LLM with legal knowledge cues, which augmented prompts with domain-specific rules (e.g., references to Secretary of State imply England, Estyn suggests Wales). Model outputs were compared against the Annex, and accuracy was computed across all 128 sections to assess performance improvements across methods. All code and evaluation tools are publicly available².

3. Results and Discussion

The baseline extent tag approach achieved 46.9% accuracy, revealing that current metadata does not reliably represent territorial scope. NER improved performance to 61.7%, while the LLM-based methods achieved 63–73% accuracy. When enriched with legal knowledge, accuracy reached nearly 80%. These results demonstrate that combining structured context and simple legal cues can significantly improve AI-based territorial inference. However, manually embedding knowledge in prompts is not scalable, emphasizing the need for integrating legal knowledge graphs [2] in future research.

4. Conclusion and Future Work

Manual identification of territorial applicability in UK legislation is slow and challenging due to nuanced legal language and cross-references. This study shows that AI models can analyse statutory text to suggest likely territorial scope, reducing dependence on manually crafted annexes. Future work will focus on exploring more techniques to add external knowledge [3] and enhancing interpretability so that each AI-generated decision is supported with clear reasoning and can be easily verified.

References

- [1] Liga D, Robaldo F. Fine-tuning GPT-3 for legal rule classification using LegalDocML and Legal-RuleML. *Artif Intell Law*. 2023.

²<https://github.com/SafiaK/legislation-territory-analyzer>

- [2] Cui J, Ning M, Li Z, Chen B, Yan Y, Li H, Ling B, Tian Y, Yuan L. Chatlaw: A multi-agent collaborative legal assistant with knowledge graph enhanced mixture-of-experts large language model. 2024.
- [3] Zhang Q, Dong J, Chen H, Zha D, Yu Z, Huang X. KnowGPT: Knowledge graph-based prompting for large language models. In: *Proceedings of the 37th Conference on Neural Information Processing Systems (NeurIPS 2024)*; 2024.