

Phelan US Centre Undergraduate Research Assistantship Report - Emma Heijnen

During my time as a Research Assistant at the Phelan US Centre, I contributed to Professor Schonhardt-Bailey's research on U.S. Public Lands. Nearly forty per cent of the 2.27 billion acres of land in the U.S. is publicly owned and managed by federal and state agencies, including the Bureau of Land Management (BLM), the U.S. Forest Service (USFS), the National Park Service (NPS), and the U.S. Fish and Wildlife Service (FWS). Although public approval for these agencies is high, and the lands themselves are central to American identity, they are nonetheless contested. There is a tension between economic development and environmental protection, and between national security and the rights of indigenous peoples.

The project attempts to measure these conflicts by examining public lands litigation across six decades, from 1960-2024. Last year's Phelan's Research Assistant laid the foundations of the project by constructing a database of federal court cases, drawing almost 5.000 relevant cases from the Westlaw legal database. The research focuses on cases across the 9th and 10th Circuit Courts, which cover the western states where the great majority of public lands are located. My work, together with three other research assistants, helped with the next phase, namely analysing and interpreting how the legal argumentation has evolved.

Given the scale of the dataset, reading and manually coding every case was not feasible. Instead, we used Large Language Models (LLMs). The LLM coded case metadata (citation, parties, circuit, data), classified litigant types (ENGOS, federal government, industry trade associations, individuals, tribal governments, and others) and identified key legal topics. The two data-science RAs handled the computational implementation, whereas Ishaan and I evaluated the LLM outputs against the original Westlaw case documents. This meant identifying inconsistencies between what the model extracted and what the case actually said, and then feeding those observations back into the prompt refinement process. I found it very rewarding to work on this interdisciplinary project where there was a continuous dialogue between the RAs and the Professor and where we all learnt from each other's (modest) expertise.

LLMs are not without limitations, particularly their tendency to 'hallucinate' is a challenge. To minimise this risk, the project used LLMs only for tasks that could not be reliably completed using traditional rule-based coding. For example, one of the data-science RAs developed a rule-based classifier using regex pattern matching to identify the winning parties. Again, our role as 'legal' RAs was to check the outputs against the original Westlaw cases and identify instances where the classifier produced incorrect results. This feedback was then directly used by the data-science RA to improve the code. Again, highlighting the importance of this interdisciplinary approach!

The most ambitious, but also most interesting, part of the project attempts to scale cases along a spectrum from 'traditional/procedural' to 'modern/right-based' argumentation. Existing legal and political science literature highlights how, before the 1970s, public lands litigation was dominated by disputes over the validity of mining claims, the terms of leases, and the authority of the Secretary of the Interior. The 1970s environmental legislation, including the National Environmental Policy Act (1970), the Endangered Species Act (1973), the Federal Land Policy and Management Act (1976), and the National Forest Management Act (1976), created procedurally challengeable governance of public lands. Legal questions shifted from whether a mining claim is valid to whether the agency correctly followed the required procedures (Environmental Assessment). Finally, more recently, there

has been a rise of ‘modern’ argumentation: environmental litigation has increasingly focused on ‘rights’, including environmental justice, rights to recreate and cultural rights of indigenous peoples.

Although these trends are documented in the literature, they are harder to capture in our database. Some of the trends are confirmed by the LLM topic analysis, for example, ‘NEPA compliance’ was one of the fastest-rising topics. However, ideally, the LLM would also code the cases along a ‘traditional, procedural, right-based’ spectrum. Unfortunately, after reviewing and testing different prompts, we determined this may not work yet because our database only includes Westlaw’s case summaries rather than the full judicial opinions. This summary, therefore, does not entirely capture the nuance of the legal arguments that the project aims to measure. To address this, Ishaan and I have complemented the LLM approach with a more ‘traditional’ qualitative manual coding of a ten per cent sample of the full dataset, working from the original, full-opinion case documents. The coding scheme we applied distinguishes between multiple dimensions, including cases primarily about federal resource administration, procedural cases reflecting an instrumental and non-instrumental view of nature, and explicitly rights-based categories covering climate change, indigenous rights and environmental justice. Many cases sit across multiple categories, which can be captured by our scoring scheme. I found it very interesting to get a more in-depth grasp of the cases through this manual coding. Additionally, it highlighted how LLMs can complement traditional research methods where a ‘human’ judgement remains indispensable.

Ultimately, I hope that the research we completed advanced the project both by refining the prompt design and by improving the understanding of how argumentative frames have shifted across the database. Both these contributions hopefully bring Professor Schonhardt-Bailey closer to answering how ‘ideas’ regarding the governance of U.S. public lands have evolved. Personally, working on this project was very formative. It was very intellectually rewarding to be exposed to this interdisciplinary research and see how legal opinions can be operationalised computationally. Professor Schonhardt-Bailey gave us a lot of independence and trusted our judgement, which was both challenging and educational. Finally, it was very interesting to learn more about the history of U.S. public lands. In my environmental law course, I could clearly see how the public lands governance has inspired regulatory frameworks worldwide, and how it exposes fundamental conflicts over natural resource extraction. Finally, I hope that the foundations laid during my time on the project will support the project and perhaps another Phelan Research Assistant next year.