

Increasing Public Data Transparency for Immigration Law in Canada

Ismail (Husain) Bhinderwala, Jessica Yu, Ke Gao, Yichun Liu

2025-05-04

Table of contents

1	Executive Summary	2
2	Introduction	2
2.1	Problem Overview	2
2.2	Why This Matters	2
2.3	Project Motivation	3
2.4	Data Sources	3
2.5	Our Solution	4
2.6	Data Science Techniques	4
3	Data science methods	5
3.1	Structured Data	5
3.2	Unstructured Text Processing & LLM Pipeline	6
4	Data Product and Results	9
4.1	Data Quality Report	9
4.2	Curated Data Stories	10
4.3	Interactive Exploratory Dashboards	11
4.4	LLM-Based Metadata Extraction Pipeline	13
4.5	Future Improvements and Alternatives	13
5	Conclusions & Recommendations	13
5.1	Recalling the Problem & Solution	13
5.2	Key Findings	14
5.3	Limitations & Recommendations	14

1 Executive Summary

This project addresses the lack of transparency and accessibility in Canadian immigration inadmissibility decisions by asking: How can data science methods uncover patterns or biases in IRCC rulings and support legal advocacy? Inadmissibility decisions have lasting effects on individuals and families, yet the underlying data is poorly documented and difficult to analyze. Partnering with Heron Law Offices, which represents clients affected by these decisions, we propose a data-driven solution. We developed a comprehensive data pipeline that includes a data quality report to assess IRCC data reliability, an LLM-based extraction system to structure information from court rulings, and curated data stories to highlight recurring themes. Our analysis applies statistical inference and natural language processing to uncover patterns and possible biases. These insights are made accessible through an interactive dashboard, enabling legal professionals to explore the data and support clients more effectively.

2 Introduction

2.1 Problem Overview

Inadmissibility¹ decisions issued by Immigration, Refugees and Citizenship Canada (IRCC) can have severe, long-lasting consequences for individuals and their families including denial of entry or removal from Canada. Yet the data underlying these decisions is often fragmented, inconsistently structured, and inaccessible to the legal professionals who most need it. This lack of transparency makes it difficult to assess systemic fairness, identify regional or racial disparities, or support legal and policy advocacy.

2.2 Why This Matters

Our project partner, **Heron Law Office**, frequently represents individuals facing inadmissibility and litigates immigration decisions in federal court. The firm has identified several concerns, including:

- The potential role of **geopolitical influence** (e.g., cases from Ukraine, Syria, China, Iran)
- Patterns suggesting **anti-Black and anti-African bias**, reflected in disproportionately high dismissal rates² from African and Caribbean countries
- Gaps and inconsistencies in the released data that limit legal interpretability and policy reform.

¹“Inadmissibility” refers to the legal grounds under IRPA that render a foreign national or permanent resident ineligible to enter or remain in Canada.

²“Dismissal rate”: the number of cases dismissed at the Leave stage divided by the total number of Leave applications.

These challenges are compounded by IRCC’s growing reliance on **Advanced Analytics** for decision-making, which further raises the stakes for transparency and oversight.

2.3 Project Motivation

Our project aims to improve data accessibility, clarity, and interpretability by applying **data science methods** to immigration-related datasets. Specifically, we translate broad legal concerns into tractable data problems, with the goal of surfacing meaningful patterns and equipping legal advocates with structured, evidence-based insights. Recognizing that most law-firm partners lack deep statistical expertise, we proactively flag potential data biases and quality issues. With their perspective in mind, we deliver a dashboard that presents only the essential visualizations, which keep our key findings clear and easy for non-technical users to understand and act upon.

2.4 Data Sources

When a person applies for a visa or permanent residency in Canada, IRCC first issues an inadmissibility decision under A34(1)³. If refused, the applicant may file a litigation application, a request for leave to judicially review that decision in Federal Court. If leave is granted, the case proceeds to a full hearing and the written court decision is published publicly.

To reconstruct this journey, we worked with three complementary datasets:

- **A34(1) Refusals Dataset (2019–2024)**: IRCC’s internal records of refusals by country, inadmissibility grounds, and status. It captures every applicant deemed inadmissible.
- **Litigation Applications Dataset (2018–2023)**: Federal Court filings detailing leave requests, case types, and decision offices. It shows who challenged their refusal and at what stage.
- **Refugee Law Lab Legal Text Dataset (2001–2024)**: The full text of Federal Court (decisions, used for our NLP analysis). This bulk open-access legal text dataset is publicly available via the Refugee Law Lab on Hugging Face “Refugee-Law-Lab / Canadian-Legal-Data - Datasets at Hugging Face” (n.d.).

By linking these sources, we can follow cases from the point of refusal through the leave application process to the ultimate court ruling. This enables end-to-end transparency and deeper insights into systemic patterns.

³A34 (1): A permanent resident or a foreign national is inadmissible on security grounds

2.5 Our Solution

To directly address the legal partner’s needs and highlight the most relevant data narratives, our proposed solution includes four integrated components:

- **Data Stories**

A curated narrative across three themes:

- Grounds for refusal by year and country.
- Regional differences in litigation outcomes.
- Country-level litigation patterns for Nigeria, India, Iran, and China.

- **Interactive Dashboard**

- *A34(1) Refusals*: Filter by year, country, and grounds.
- *Litigation Applications*: Navigate trends by case type, leave decision, and region.

- **LLM-Based Pipeline**

A legal text processing pipeline using **LLaMA-3** and regular expressions to extract:

- Judge names, decision outcomes, cities.
- Case classification (e.g., inadmissibility vs. others).

- **Data Quality Report**

A structured evaluation identifying:

- Missing disaggregations (e.g., country-level variables in refusals).
- Inconsistencies in naming, case categorization, and court metadata.
- Recommendations for improving future transparency and open data releases.

2.6 Data Science Techniques

To meet our goals of improving data clarity, uncovering systemic patterns, and empowering legal advocates, we use the following data-science techniques to address specific objectives.

- To surface meaningful patterns in litigation application, we conducted **exploratory and inferential analysis**, using chi-square tests on contingency tables of litigation counts by country and case type. Chi-square is ideal here because our data are purely categorical counts with sufficient expected cell frequencies, which allow us to detect statistically significant differences in outcome distributions across groups.
- To ensure data quality, reproducibility, and ongoing transparency, we developed modular pipelines for **data cleaning**, which automatically normalize outcome codes and impute missing values, and for **legal text processing**.

- To make insights accessible for legal advocates, we built an advocacy-oriented dashboard aligned with **Heron Law’s public interest goals**. The dashboard features guided Data Stories, contextual annotations for terminology and caveats, and a free-exploration module so users can uncover patterns in immigration decisions and support more equitable, transparent analysis.

Together, they support legal professionals, researchers, and policymakers in better understanding Canadian immigration decisions and in advocating for the fairness and openness that the system demands.

3 Data science methods

This section outlines our complete data science workflow, from initial exploration to final deployment. Our methodological decisions were motivated not only by technical feasibility, but also by client needs, data limitations, and legal-ethical considerations. We approached this work iteratively, continuously refining methods and outputs in response to stakeholder feedback.

3.1 Structured Data

3.1.0.1 Phase 1: Data Cleaning & Preprocessing for IRCC Datasets

We began by transforming the raw A34(1) refusal records into a tidy format with one row per applicant and refusal ground. We standardized country names and outcome codes to eliminate inconsistencies, for example treating “Democratic Rep. of Congo” and “Congo, Democratic Republic of the” as the same entity, and we systematically identified and imputed missing values using rule-based logic. Because A34(1) refusals were sparse and unevenly distributed, with most countries reporting fewer than 50 cases across all years, we prioritized exploratory analysis over formal statistical inference. These steps established a consistent and reliable dataset that ready to use for next step.

3.1.0.2 Phase 2: Exploratory Data Analysis & Story Development

In the first four weeks, we conducted **exploratory data analysis (EDA)** to identify emergent patterns and generate hypotheses, revealing a sharp rise in Ukrainian inadmissibility refusals in 2024 (from 5 to 134 cases), a similar increase for Syria in 2022 (from 9 to 46 cases), and a marked uptick in **mandamus**⁴ litigation by Chinese applicants in 2023. We validated these findings with our partner to ensure they aligned with court-level observations, and from that discussion we prioritized data stories on **Geopolitical shifts** (Ukraine, Syria

⁴Mandamus a judicial writ issued as a command to an inferior court or ordering a person to perform a public or statutory duty.

and China), **Anti-Black/African bias** (notably high dismissal rates^[13] for African and Caribbean countries) and **gaps in disaggregated data** by office or demographic attributes. We then iteratively refined our visualizations by applying principles from SAGE Research Methods (2025) and Storytelling With Data (2025). We transformed generic chart titles into narrative-driven insights. For example, “Refusal Counts by Country” became “Ukrainian A34(1) Refusals Jump from 5 to 134 Cases (2019–2024)”, “Dismissal Rate by Continent” was reframed as “African and Caribbean Applicants See 15–20 pp Higher Dismissal Rates vs. Global Average”.

3.1.0.3 Phase 3: Statistical Analysis

To determine whether observed patterns varied significantly by country, we conducted statistical analysis and applied chi-square tests, which are ideal for categorical comparisons without distributional assumptions. While logistic regression could provide finer-grained estimates, we avoided it due to low sample sizes and missing covariates. The disadvantages are that it cannot adjust for confounders and sensitive to small cell counts. Two hypotheses were tested for the top 4 countries (Nigeria, Iran, India, China): the **distribution of case types** is independent of country and that **distribution of decision outcomes** is independent of country. Both were rejected ($p < 2.2 \times 10^{-1}$), indicating that applicant nationality is associated with different litigation experiences. However, these results are descriptive, not causal. They apply only to countries with sufficient sample sizes and cannot control for confounders due to the lack of joined demographic data.

3.1.0.4 Phase 4: Visualization & Dashboard Design

We developed an interactive dashboard with Streamlit and Plotly that guides users through **three data story pages** (A34(1) Refusals, Litigation Outcomes, Country-Level Trends) and **two interactive explorers** (Litigation and A34 (1) Refusals). To ensure accessibility for non-technical users and clear visual narratives, we embedded inline tooltips, preset filters, and in-chart caveats. Our unified interface delivers both curated insights and the freedom to dive deeper. We applied a sequential BuGn color palette from ColorBrewer Brewer (2025), adjusting the number of classes per chart. The limitation is that the scheme may not be fully color-blind friendly.

3.2 Unstructured Text Processing & LLM Pipeline

To process Canadian Federal Court decisions (2014–2024) curated by the Refugee Law Lab, we implemented a multi-stage pipeline that combines **rule-based methods** (Regex) with **large language models (LLMs)** for zero-shot classification and attribute extraction. This hybrid approach was necessary to handle the linguistic complexity and semi-structured nature of legal decisions. As shown in Figure 1, the pipeline consists of multiple stages, integrating both deterministic and AI-driven components.

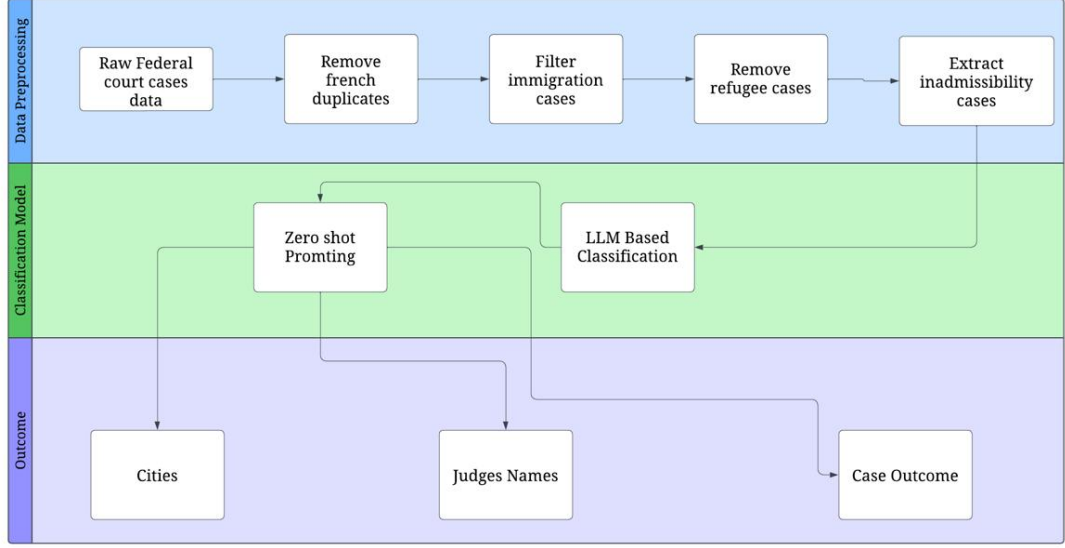


Figure 1: LLM-based multi-stage pipeline for legal decision processing.

3.2.0.1 Phase 1: Preprocessing & Filtering

We first restricted our **Temporal scope** to Federal Court decisions issued between 2014 and 2024. We then **normalized language** by removing duplicate French texts whenever an English equivalent could be identified via standardized citation strings. Next, we applied **topical filters** with using regular expressions to include only cases that mention immigration-related authorities (e.g. “Citizenship and Immigration” or “MCI”) and to exclude refugee protection claims. Finally, we **focused on inadmissibility** by isolating decisions that contain terms such as “inadmissible” or “inadmissibility.”

3.2.0.2 Phase 2: Inadmissibility Classification

At this phase, we used **regular expressions** to detect explicit IRPA citations such as “s.34” or “section 36.” When decisions lacked these direct references, we applied semantic fallback patterns by looking for terms such as “espionage” or “indictable offence” to infer the relevant statutes. Any cases that remained ambiguous or unclassified were grouped under an “other” category.

3.2.0.3 Phase 3: LLM-Based Extraction & Validation

We deployed LLaMA 3 via Ollama on a local server to extract decision-level metadata using targeted prompts Rehaag (2023). We extracted judge names from the top 30 lines of each decision, parsed the hearing city from lines 10 to 25, and retrieved the decision outcome from

the final 20 to 50 lines. We adopted this approach instead of traditional NLP pipelines, such as spaCy combined with rule-based parsers, because those methods struggled with the variability of legal prose and demanded extensive manual rule creation, whereas LLMs offered stronger generalizability for under-structured texts. To address the risks of non-determinism and hallucinations inherent in LLM outputs, we implemented prompt tuning and followed every automated extraction with manual validation. Future improvements might include fine-tuning a domain-specific LLM or employing ensemble prompting strategies to boost robustness, though these would require additional compute resources and labeled data beyond this project’s scope.

3.2.0.4 Phase 4: Manual Review & Reproducibility

At the final phase, we asked our partner to manually validate a sample of cases, and they reported approximately 85 percent accuracy for both classification and attribute extraction. They also identified several challenges: LLM outputs varied unless sampling parameters were fixed; small changes in prompt phrasing sometimes produced different results, and differences in hardware affected both consistency and processing time. Despite these issues, the LLM pipeline substantially outperformed purely rule-based methods and proved essential for scaling the extraction of structured insights from noisy legal text. Rehaag (2023)

3.2.1 Evaluation Criteria and Metrics

Component	Metric / Method	Notes
Statistical Testing	p -value from Chi-Square Tests	Used only where assumptions were met
LLM Pipeline	Accuracy (Manual Validation)	Partner-verified; ~85% on early sample
Data Quality	Missing Value Rate, Consistency Checks	Shared as part of data quality report

3.2.2 Methodological Limitations & Assumptions

- **Data Sparsity:** Many variables (e.g., refusals by country-year-ground) had too few entries (<5) for valid statistical testing.
- **No Causal Inference:** All findings are observational. Due to missing contextual variables, **no claims of discrimination or intent** can be robustly made.
- **Unjoined Datasets:** Key datasets (e.g., litigation + refusal) could not be linked at case level, limiting multivariate analysis.
- **Office-level Analysis:** More than half of cases lacked decision office information, making regional bias analysis incomplete.

- **LLM Variability:** Extraction outputs varied by run; reproducibility was partially mitigated through seeding and manual verification.

3.2.3 Alternative Products Considered

- A **Shiny app** could offer finer-grained UI, but would increase deployment complexity and cost.
- **Fully supervised LLM classification** for decision outcomes would improve accuracy, but requires hand-labelled data and compute resources beyond current scope.

3.2.4 Stakeholder & Ethical Considerations

Our work supports immigration lawyers, legal advocacy groups, and policymakers, but it also carries ethical risks such as **overinterpretation of sparse data**, **biases introduced by LLM extraction**, and potential **privacy concerns from sensitive court text fields**. To mitigate these risks, we explicitly documented all data limitations, embedded caveats and disclaimers within the dashboard, and refrained from drawing conclusions when statistical evidence was insufficient.

4 Data Product and Results

Our data product was designed to address the dual priorities of our legal advocacy partner: (1) identifying trends in inadmissibility and litigation that may reflect systemic bias or inconsistencies, and (2) enabling future research and policy work grounded in high-quality legal data. To meet these goals, we delivered an integrated data product composed of: a **Data Quality Report**, **Curated Data Stories**, **Interactive Exploratory Dashboards**, and a **LLM-based Pipeline** for extracting metadata from unstructured legal decisions. These components work together to create a reusable foundation for empirical legal research, evidence-based advocacy, and internal capacity-building for legal professionals.

4.1 Data Quality Report

The **Data Quality Report** was developed to provide transparency into the structure and limitations of the two IRCC datasets, A34(1) Inadmissibility Refusals and Litigation Application Decisions. Rather than treating the datasets as complete, we identified critical gaps so the partner could better interpret patterns and advocate for improved data transparency from IRCC. This report is intended to help the partner challenge the evidentiary value of government data when it is incomplete or inconsistently recorded. For instance, undocumented

decision offices or missing filing dates directly affect the ability to analyze regional bias or litigation timelines.

Examples of Key Gaps

- Most primary decision-office fields were missing, limiting regional analysis.
- A34(1) refusal data lacked case counts for permanent residents in several years.
- The litigation dataset omitted filing dates restricting temporal and procedural analysis.

While we cannot fix these omissions, documenting them is itself valuable: it allows legal stakeholders to contextualize quantitative findings and identify where further qualitative or legal review is needed. However, IRCC may not acknowledge or rectify these issues, so our product avoids overpromising and instead offers cautious interpretation of patterns based on known limitations.

4.2 Curated Data Stories

We designed three static and narrative-style data stories that present insights in a structured and accessible way. These stories were developed in collaboration with our partner and validated through regular review meetings to ensure they reflect legal relevance. Curated stories serve advocacy needs more effectively than open-ended plots. They are shareable, interpretable, and focused on narratives the client already cares about (e.g., potential systemic bias, country-specific litigation trends). They also help prevent overinterpretation of weak signals in the data, especially given the quality issues flagged in our report. We also added concise, step-by-step “how to read this chart” captions to each figure to guide our nontechnical users through the visualizations.

Below are a few polished example graphs from those data stories. To explore the full interactive dashboards, please visit our GitHub repository: <https://github.com/ismailbhinder/Public-Transparency-for-Immigration-Law-Canada.git>. For detailed local setup and deployment instructions, see the README.md in the repo’s root directory.

Story Overviews & Client Relevance

1. A34(1) Refusals by Country and Ground

- *Insight:* A34(1)(f)⁵, was the most cited inadmissibility ground, with Ukraine and Syria showing major spikes as seen in Figure 2
- *Why it matters:* Highlights geopolitical patterns that may warrant further scrutiny.

2. Litigation Outcomes by Region and Country

⁵Being a member of an organization that is believed to be involved in espionage, subversion, or terrorism.

Unable to display output for mime type(s): text/html

(a) Heatmap of A34(1) refusals by country and year (Top 5 countries).

Unable to display output for mime type(s): text/html

(b)

Figure 2

- *Insight*: Nigeria dominated RAD⁶ appeals; Iran had a spike in visa refusals post-2020 as shown in Figure 3
- *Why it matters*: Supports country-specific advocacy or policy submissions based on litigation profiles.

Unable to display output for mime type(s): text/html

Figure 3: Litigation outcomes by case type and country for selected regions (China, India, Iran, Nigeria).

3. Litigation Trends Over Time

- *Insight*: Steady increase in federal litigation since 2018, dip in 2020 (COVID), then strong rebound as seen in Figure 4
- *Why it matters*: May reflect broader procedural shifts, backlogs, or access-to-justice barriers.

Unable to display output for mime type(s): text/html

Figure 4: Trend in federal court litigation counts by year.

4.3 Interactive Exploratory Dashboards

To complement the curated stories, we built two interactive dashboards using Streamlit and Plotly. These tools allow legal professionals or researchers to filter by case type, country, refusal ground, or time range, revealing tailored views that support investigative or legal work.

Use Case & Justification

⁶RAD is the Refugee Appeal Division of the Immigration and Refugee Board of Canada, which hears appeals of first-instance refugee decisions.

While curated stories convey known patterns, exploration empowers discovery, especially for legal practitioners preparing country reports, litigation strategies, or academic research. Interactivity is essential for users who wish to dig deeper into their own areas of interest.

Examples

- A lawyer representing a Syrian client could use the A34(1) explorer to review trends in inadmissibility refusals.
- A researcher could filter the litigation dashboard to compare decision outcomes for Iranian vs. Indian applicants across years.

Design Priorities

- Clean layout with minimal cognitive load: key metrics (total litigation count, countries selected, year range) surface in cards up top while filters live in a collapsible sidebar (see Fig Figure 5).
- Hover tooltips and annotations for interpretability: mousing over map regions or bars reveals definitions (e.g., “LIT Litigation Count”) and flags low-volume cells
- Filters for key attributes (country, year, ground, etc.): preset defaults for country, year, and case type prevent missteps, with “Clear All Filters” for full exploration.

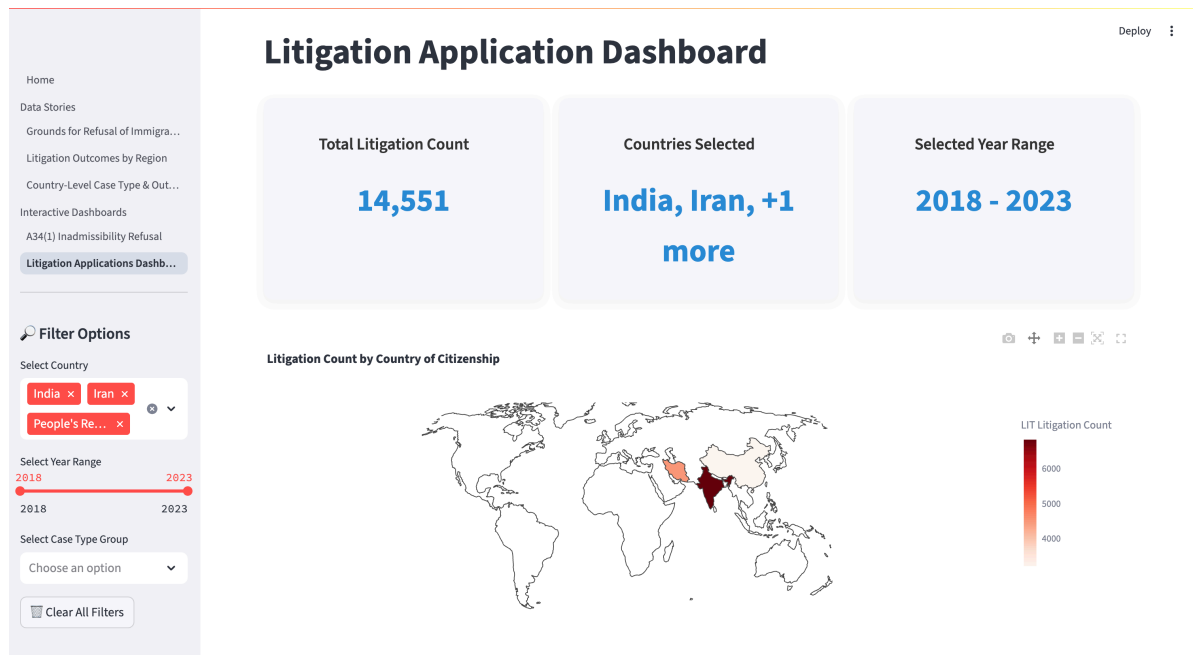


Figure 5: A snapshot of the Interactive Dashboard

Limitations & Risks

While exploration is valuable, it can also lead to misinterpretation of noisy or sparse data. To

mitigate this, we embedded tooltips and footnotes throughout the dashboard to warn users of known data quality concerns. We also advise our partner to treat these tools as starting points for inquiry, not sources of final proof.

4.4 LLM-Based Metadata Extraction Pipeline

As a final component, we built a local LLM pipeline using LLaMA 3 (via Ollama) to extract structured metadata from unstructured court decisions—supporting long-term capacity building by allowing the partner to re-run the workflow whenever new decisions are released. The same pipeline can filter and extract inadmissibility cases across any year and is adaptable to Federal Court, Federal Court of Appeal, or Supreme Court decisions by simply adjusting line numbers or prompt templates. We chose LLMs because manual extraction isn’t scalable and rule-based approaches can’t handle legal-language complexity, and our pipeline reliably pulls out key attributes like decision outcome, judge name, and hearing location even from unstructured text. Its advantages include timely updates to litigation datasets, ~85 % accuracy in partner validation, and generalizability across years and document formats. Its drawbacks are non-determinism without careful prompt engineering and seeding, hardware-dependent reproducibility, and occasional failures in edge cases (e.g., multiple-judge panels or ambiguous outcomes).

4.5 Future Improvements and Alternatives

In future iterations, we plan to enhance the dashboard with automated summaries and trend annotations that surface key shifts without manual interpretation, add hoverable definitions for complex legal terms such as “mandamus” or “RAD decisions” to support non-experts, and refine our LLM prompt design by enforcing explicit output templates and structured responses to reduce nondeterminism and prompt sensitivity in metadata extraction.

5 Conclusions & Recommendations

5.1 Recalling the Problem & Solution

Our partner challenged us with this question: *How can data science tools uncover patterns in Canadian inadmissibility* and litigation to help legal practitioners, advocates, and policymakers promote transparency and fairness? To meet that need, we built a layered data product combining a **Data Quality Report**, **Curated Data Stories**, **Interactive Exploratory Dashboards**, and an **LLM-based extraction pipeline**. At the **case level**, the LLM pipeline transforms unstructured Federal Court decisions into structured metadata (outcome, judge, location), enabling broad pattern analysis. Then curated stories surface geopolitical shifts (e.g., Ukraine, Syria) and systemic disparities in refusal and litigation rates;

and the data quality report highlights gaps (e.g., missing office or refusal-ground details) that can inform FOI requests and policy dialogue with IRCC. This integrated approach balances narrative clarity with exploratory flexibility and was refined continuously through partner feedback.

5.2 Key Findings

- **Inadmissibility** refusals showed major spikes for Ukraine (2024) and Syria (2022), primarily under s.34(1)(f), with higher rates for permanent residents.
- **Litigation case volumes** rose significantly post 2021, with country-specific trends, Nigeria in RAD appeals, Iran in visa refusal challenges.
- **A34(1)** refusals and litigation outcomes differ by country, supported by chi-square tests, though confounding factors cannot be ruled out.
- **Legal text extraction** using LLMs achieved ~85% accuracy in pilot tests, enabling scalable, semi-automated analysis.

5.3 Limitations & Recommendations

Our work faces several key limitations: the LLM’s classification and extraction remain probabilistic and require expert legal review to catch edge-case errors. There’s critical data quality gaps, including undefined field meanings, missing application dates, absent demographic co-variables and reliance on aggregate counts prevent detailed country level trend analysis. Because litigation and refusal records cannot be joined at the case level, we are unable to model decision pathways or systemic bias robustly.

Future iterations of the LLM pipeline should embed **human-in-the-loop validation** to ensure expert review of extracted features and minimize error propagation. We will integrate **advanced statistical modules**, such as negative-binomial and zero-inflated models, to better model count data and rare events. Our findings should inform targeted data advocacy, urging IRCC to publish more complete and **structured fields** such as approval rates, decision authority, and refusal reasoning. Finally, we will **uphold sustainability practices** by maintaining reproducible environments and comprehensive prompt logs to guarantee transparency as the project scales or transitions to new teams.

5.3.1 Final Reflection

This project demonstrates how data science can meaningfully support legal transparency and advocacy but also reveals the fragility of analysis when working with **fragmented public data**. Our work provides a **prototype for responsible, reproducible legal data analysis**, and underscores the need for **ongoing collaboration between data scientists and legal experts** to ensure interpretations remain grounded, ethical, and actionable.

-
- Brewer, Cynthia A. 2025. “ColorBrewer2 Sequential Palette (BuGn, 3 Classes).” <https://colorbrewer2.org/#type=sequential&scheme=BuGn&n=3>.
- “Refugee-Law-Lab / Canadian-Legal-Data - Datasets at Hugging Face.” n.d. <https://huggingface.co/datasets/refugee-law-lab/canadian-legal-data>.
- Rehaag, Sean. 2023. “Luck of the Draw III: Using AI to Extract Data about Decision-Making in Federal Court Stays of Removal.” *Queen’s LJ* 49: 73.
- SAGE Research Methods. 2025. “Telling a Data Story: How to Develop a Compelling Data-Driven Narrative.” <https://methods.sagepub.com/how-to-guide/telling-a-data-story-how-to-develop-a-compelling-data-driven-narrative/>.
- Storytelling With Data. 2025. “Makeovers.” <https://www.storytellingwithdata.com/makeovers>.