CSC490 Assignment A1 - Project Landscape

Armaan Rehman Shah - 100941309
Boaz Cheung - 1007673607
Alice Sedgwick - 1009301355
Yuan Yu - 1008782195

September 14, 2025

1 Part One: Interest Statements

Immigrating is notoriously difficult with language barriers, fragmented information sources, and expensive attorney fees. The project we would like to work on is an international immigration documentation assistant website. It would help people immigrating to various nations better understand immigration requirements, prepare their documents, and reduce costly delays caused by rejected submissions. As a team, we are passionate about improving immigrant lives and eager to apply modern approaches such as retrieval-augmented generation (RAG) to make the immigration journey more accessible and less overwhelming.

Individual Comments:

- Boaz - As an immigrant myself, I understand first-hand the complexity of the many immigration forms, and it often feels overwhelming to be without anyone who can answer my questions.

- Armaan - I remember how frustrating it was to apply and renew my Canadian study permit, especially with the IRCC website full of broken links. I am excited to work on the RAG pipeline to make immigration information clearer and more accessible.

- Alice - Hearing all the immigration horror stories, I cannot help feeling excited to work on this project to help immigrants, and I hope to improve my ML development skills and teamwork along the way.

- Yuan - I have been through the application process of Canadian Visa, study permit, and Co-op permit. It is indeed truly hard to navigate through the bloated IRCC website.

2 Part Two: Landscape Analysis

Table 1: 7 Relevant companies, 2 papers, and 2 open source projects

Relevant Item	Description	Commentary
CBR-RAG (Case-Based Reasoning for RAG in LLMs for Legal QA)	A paper that proposes combining Case-Based Reasoning (CBR) with Retrieval-Augmented Generation (RAG) to improve legal question answering.	Strength is in structured retrieval (cases as reusable knowledge). Use this technique in the immigration field instead of the law where this paper was originally based on..
Building a Robust Retrieval Augmented Generation Chatbot for Immigration Knowledge Base	AI immigration chatbot built on Google Cloud components especially Vertex AI for embeddings and Gemini for LLM	A paper that is directly relevant containing RAG and data pipelines. Still contains some points for improvements such as chained reasoning and time-aware retrieval since immigration rules change very rapidly.
Parley	Automates research, drafting, organization of evidence for US employment visas like EB-1, O-1 etc.	An example of a US-market focused tool. Assist law firms scale by handling document and evidence assembly, freeing up professional time.
Lisa AI	Offers AI powered US visa guidance, document analysis, and helps users connect with immigration attorneys and advisors	Contains a Marketplace + AI agent hybrid model. Which blends DIY with expert knowledge.
Pathpr	Canadian tool to assess profile, matching to immigration programs, PNP guidance, Express Entry estimates.	Feature to match users with their best fit immigration pathway. Focus on eligibility of programs instead of application process.
VusaVit	Visa requirement databases, online application/processing, appointment booking, mock interview tools.	An example of productized visa-requirements databases and automation for appointment/processing flows. Doesn't contain that much LLM features
OpenSphere.ai	AI-powered immigration assistance, legal guidance, visa approval maximization for US work/investor visas	Contains an AI application review feature and also offers a community feature were similar applicants are able share insights with each other
Tere.app	AI agent to answer immigration queries about the Spanish Immigration system	An example of localized advice for Spain, however lacks English or other lanuage support.
ChatMigrante	Chat agent for people get information about the US immigration in their native langauge	Excellent example of language-tailored product, bridging information gaps and serving vulnerable populations.
ibrotzky/visabot	Open-source immigration chatbot where you can ask questions about the canadian immigration process	Uses hard coded requirements and rules instead of an LLM, shows a more dated approach.
immigration-ai	Open-source chat interface using FastAPI and OpenAI's API for U.S. skilled immigration questions.	Contains embedding search to query for immigration documents. Contains reliance on OpenAI but a good application baseline.

3 Part Three: Project Outline

3.1 Problem Statement

Immigration processes are fragmented across countries, with each government publishing distinct rules, forms, and eligibility requirements on their own websites. Applicants face three recurring challenges: (1) rules scattered across multiple sources, (2) legal language that is difficult to interpret, and (3) lack of a unified platform for reliable, country-specific answers. This results in confusion, wasted time, and frequent mistakes in applications for study, work, or permanent residency.

3.2 Proposed Solution

We propose building a robust immigration-focused chatbot powered by Retrieval-Augmented Generation (RAG) and a fine-tuned Large Language Model (LLM). Unlike generic assistants, our system is designed to operate as a reliable, transparent tool for navigating multiple countries' immigration rules. Specifically, the chatbot will:

- integrate and maintain immigration content from 3–4 official government websites (e.g., Canada IRCC, U.S. USCIS, UK Home Office, Australia DHA)

- preserve country-specific distinctions while consolidating information into a single access point

- generate source-cited responses that directly link back to official text, ensuring reliability and trust

- provide a unified, intuitive interface for applicants navigating different immigration systems.

3.3 High-Level Technical Approach

Our four-month roadmap is:

1. **Data Preparation:** Scrape and preprocess immigration content; store embeddings in a vector database (e.g., FAISS or Pinecone) with country and visa-type tags.

2. **Prototype Pipeline:** Connect a base LLM (Llama 3, Mistral, or GPT-4 API) to the database; implement a retrieval layer over the multi-country vector database. Responses will be formatted with explicit country labels and inline citations pointing back to official sources.

3. **Refinement:** Use prompt engineering or lightweight fine-tuning for clarity and neutrality; add a country-classification mechanism. Add consistency checks so conflicting information across countries is clearly distinguished

4. **Integration & Demo:** Build a simple interface (Streamlit or React), test with sample queries, and deliver a working prototype.

3.4 Key Unknowns

Important uncertainties include:

- **Data access:** Are government sites easily scrapeable or subject to legal/API limits?

- **Scalability:** Can one RAG pipeline handle overlapping corpora across countries?

- **Response Clarity:** How to ensure the chatbot explicitly distinguishes between conflicting rules when user queries are ambiguous?

- **Evaluation:** Should accuracy be judged by source match, expert review, or user comprehension?

- **Maintenance & Risk:** How to keep rules updated and ensure users do not misinterpret outputs as legal advice?

4 Part Four: Project Press Release

Immigreat: The International Immigration Documentation Assistant to Simplify Immigration for Millions
Date: September 14, 2025

Toronto, ON – Today, a new web platform called **Immigreat** was announced, offering immigrants a smarter, simpler way to prepare immigration documents and navigate the complex requirements of moving to a new country.

Immigration has long been plagued by language barriers, scattered information sources, and costly lawyer fees. Immigreat tackles these problems head-on by combining user-friendly design with cutting-edge AI to ensure that applicants understand requirements, complete their documents correctly, and avoid expensive delays from rejected submissions.

"Immigration should not be a process to dread," said **Boaz Cohen**, Co-Founder. *"It is clear that the barriers to immigration go beyond those experienced after arrival, with language barriers, fractured information, and endless confusing documents inhibiting people from reaching their new home. We built Immigreat so that nobody has to feel alone or confused during this life-changing journey."*

Key Features of Immigreat

- Step-by-step guidance tailored to the immigration program and destination country.

- AI-powered documentation assistance to flag missing or inconsistent information before submission and suggest helpful edits.

- Multilingual support that explains requirements in your native language.

- Personalized recommendations for resources, local services, and next steps.

Early Impact

"I saved over $2,000 in legal fees and cut weeks off my application process thanks to Immigreat," said **Jane Smith**, a nurse moving to Canada from Brazil.

"This is the kind of tool our settlement agencies have been waiting for. It reduces confusion, saves staff time, and makes immigrants feel empowered," said **David Lam**, Director at a newcomer services non-profit.

About Immigreat

The team behind Immigreat is motivated by a simple belief: immigration should be a great adventure, not an obstacle course. By making information clearer and the process less intimidating, the platform aims to improve lives at a global scale.

Immigreat will be available in beta later this year, with initial support for Canada, the U.S., and Japan, and expansion to other destinations planned for 2026.

Appendix

Iteration 1

Documove: The Canadian Immigration Documentation Assistant to Simplify Immigration Paperwork

Date: September 14, 2025

Toronto, ON – Today, a new web platform called **Documove** was announced, offering immigrants a smarter, simpler way to prepare their Canadian immigration documents and navigate the complex requirements of applying to Canada.

Canadian immigration is often described as overwhelming: confusing requirements, language barriers, and costly legal consultations leave many applicants stressed and uncertain. Documove addresses these challenges by combining user-friendly design with advanced AI tools that help applicants understand requirements, prepare their paperwork accurately, and avoid costly delays caused by rejected submissions.

"Immigration should not feel like a gamble," said **Boaz Cohen**, Co-Founder. *"As someone who has navigated Canadian immigration myself, I know the frustration of unclear forms and scattered information. We built Documove so newcomers can approach this process with clarity and confidence."*

Key Features of Documove

- Step-by-step guidance tailored to Canadian immigration programs such as Express Entry, study permits, and family sponsorship.

- AI-powered documentation checks to flag missing or inconsistent information before submission.

- Multilingual support that explains Canadian immigration requirements in plain, accessible language.

- Personalized recommendations for next steps and settlement resources in Canada.

Early Impact

"I was lost trying to figure out Express Entry on my own. Documove broke it down into simple steps and saved me weeks of frustration," said **Eliza Leblanc**, a nurse moving to Toronto from Mexico.

"This tool will be a game-changer for newcomers. It saves time for both applicants and the agencies that support them," said **David Lam**, Director at a Toronto newcomer services non-profit.

About Documove

Immigration is a life-changing experience that shouldn't be delayed by lengthy administrative processes. The Documove team is dedicated to making your dreams a reality with minimal pain.

Documove will be available in beta later this year, supporting the most common Canadian immigration streams, with expansion to additional pathways planned in 2026.