

Blindspot Labs: Dublin Planning AI Assistant

The Strange Data Project — Nomad AI Competition 2025

1. How It Works

Blindspot Labs is a Retrieval-Augmented Generation (RAG) system that gives LLMs access to thousands of Dublin City Council planning application records — information that no baseline model has ever seen. The system enables accurate, grounded answers to questions about specific planning applications, decisions, locations, and trends across Dublin city.

System Pipeline:

Stage	Implementation
Data Source	ArcGIS Irish Planning Applications API (Dept. of Housing) — public, no auth
Download	Python requests paginates the API, filtering for Dublin City Council
Processing	Dates converted from epoch ms, decisions normalised, appeals merged
Chunking	Each application becomes a structured text document with all fields
Embedding	OpenAI text-embedding-3-small generates vector embeddings
Storage	ChromaDB persistent local vector database
Retrieval	Top-10 semantic search on user queries
Generation	GPT-4o-mini (or Claude) produces grounded answers citing planning refs
Interface	Streamlit chat UI with optional source attribution

Model choice: GPT-4o-mini for generation (fast, cost-effective) with OpenAI text-embedding-3-small for embeddings. Claude is also supported as an alternative generator. No fine-tuning or retraining — all enhancement comes from retrieval.

2. How We Got the Data

Source: The Department of Housing publishes a national Irish Planning Applications dataset via a public ArcGIS Feature Service. The API provides structured records for every planning authority in Ireland, updated regularly, with no authentication required. We filter for Dublin City Council and pull thousands of records with full details.

Why this data improves performance: No LLM has been trained on this database. When asked about specific planning references, recent applications in a given area, or whether a particular development was granted permission, baseline models either refuse to answer, provide generic planning advice, or hallucinate fake reference numbers. This data is *functionally invisible* to LLMs — it sits behind an ArcGIS API in structured JSON that was never crawled for training.

Fields collected: ApplicationNumber, DevelopmentAddress, DevelopmentDescription, ApplicationType, ApplicationStatus, Decision, ReceivedDate, DecisionDate, GrantDate, ExpiryDate, AppealRefNumber, AppealStatus, AppealDecision, NumResidentialUnits, FloorArea, and links to full application files. Data covers 2016 to present.

Processing: Records are fetched via paginated API calls (2 000 per page), dates converted from epoch milliseconds, decisions and appeal data normalised, and each application converted into a structured text

document. Documents are embedded and indexed in ChromaDB for sub-second retrieval.

3. Real-World Application

Planning permission data is critical for multiple professions in Ireland, yet extremely difficult to query. The current system requires users to navigate the Agile Applications portal, search by individual reference number or keyword, and manually read through results. There is no way to ask natural-language questions like "what extensions were refused in my area?" or "what developments are planned near this address?"

Target users: Architects researching precedent before submitting applications. Solicitors conducting due diligence on property transactions. Property developers assessing areas for investment. Estate agents informing buyers about nearby developments. Homeowners checking what has been approved or refused near their property. Journalists investigating development trends. All of these users currently spend hours on council portals. Our system answers their questions in seconds.

4. Future Potential

This system currently covers Dublin City Council, but the ArcGIS API contains records for **every local authority in Ireland**. Expanding to full national coverage requires changing one API filter parameter. The same architecture applies to planning data across the UK (where similar open-data portals exist) and the EU.

The commercial path is clear: a SaaS product for property professionals that combines planning application data with zoning maps, development-plan policies, and property-registry information into a single AI-powered research tool. The Irish property and construction sector is worth over EUR 30 billion annually — a tool that saves architects, solicitors, and developers hours of manual research per project has obvious product-market fit. This prototype demonstrates the core value proposition: give an LLM the right data, and it becomes a planning expert.

5. How to Run

Requirements: Python 3.10+, an OpenAI API key (for embeddings and chat generation). Optionally an Anthropic API key can be used for the generation step.

Step	Command
1. Install dependencies	<code>pip install -r requirements.txt</code>
2. Configure API key	Copy <code>.env.example</code> to <code>.env</code> , add <code>OPENAI_API_KEY</code>
3. Download data + build DB	<code>python download_data.py</code>
4. Launch chat interface	<code>streamlit run app.py</code>
5. Open in browser	<code>http://localhost:8501</code>

Step 3 takes approximately 5-10 minutes. It downloads all Dublin City Council planning records from a public ArcGIS API (no auth required), processes them, generates embeddings, and builds the local ChromaDB vector database. After that the Streamlit chat interface is ready to use. No GPU or special hardware required.

6. Sample Prompts to Test

Copy and paste these prompts into the chat interface. For comparison, try the same prompts in ChatGPT, Claude, or Gemini to see the difference.

1. What planning applications were submitted in Drumcondra?
2. Was planning permission granted for developments on Griffith Avenue?
3. Show me planning decisions that were refused in Dublin 8
4. What types of residential developments have been proposed in the Docklands area?
5. Are there any appeals lodged for planning applications in Rathmines?
6. What planning applications involve demolition in Dublin city centre?

Expected results: Baseline models will either refuse ("I don't have access to planning databases"), give generic advice ("You should check your local planning authority"), or hallucinate fake reference numbers. Our system returns actual planning references, real addresses, correct decisions, and accurate dates — because it has the data.