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 20+ years of Dublin City Council planning application data — 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 required
Download	Python requests library paginates through API, filtering for Dublin City Council records
Processing	Dates converted from epoch ms, decisions normalized, appeals merged into application records
Chunking	Each application becomes a structured text document with all fields
Embedding	OpenAI text-embedding-3-small generates vector embeddings
Storage	ChromaDB persistent vector database for fast similarity search
Retrieval	Top-10 semantic search against user queries
Generation	GPT-4o-mini (or Claude) generates grounded answers citing specific planning references
Interface	Streamlit chat UI with source attribution

Model choice: We use GPT-4o-mini for generation (fast, cost-effective) with OpenAI text-embedding-3-small for embeddings. The system also supports Claude as an alternative generator. No fine-tuning or retraining is involved — 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. This API provides structured records for every planning authority in Ireland, updated regularly, and requires no authentication. We filter for Dublin City Council records only, pulling thousands of planning applications with full details.

Why this data improves performance: No LLM has been trained on the Irish Planning Applications database. When asked about specific planning references, recent applications in a given area, or whether a particular development was granted permission, baseline models (ChatGPT, Claude, Gemini) 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, is updated regularly, and exists in structured JSON format that was never crawled for training.

Data files collected:

The ArcGIS API returns rich records with: ApplicationNumber, DevelopmentAddress, DevelopmentDescription, ApplicationType, ApplicationStatus, Decision, ReceivedDate, DecisionDate, GrantDate, ExpiryDate, AppealRefNumber, AppealStatus, AppealDecision, NumResidentialUnits, FloorArea, and direct links to the full application files. Data covers 2016 to present.

Processing: Records are fetched via paginated API calls (2000 per page), dates are converted from epoch milliseconds to readable format, decisions and appeal data are normalized, and each application is converted into a structured text document optimized for semantic search. The 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 it is 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 upcoming developments. Homeowners checking what has been approved or refused near their property. Journalists investigating development trends. All of these users currently spend hours manually searching 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 only changing one API filter parameter. Beyond that, 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. 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.