



The library 2.0 – all of Switzerland's data through one gate, and the pulse running through it.

THE DATA

- + 8.7M documents · 223 sources · 4+ languages (DE/FR/IT/RM)
- + Courts · federal + cantonal law · parliaments · open-access research
- + 4.3M full texts · 8.1M embedded documents · one unified index

THE TECH STACK

- + Python 3.12 · FastAPI · OpenSearch (hybrid BM25 + kNN)
- + Qwen3-8B embeddings (1024-dim) + cross-encoder reranking
- + PostgreSQL · ClickHouse · SvelteKit · 4-tool MCP server
- + Open source (EUPL-1.2), running on UZH ScienceCloud

EXTENSIBLE BY DESIGN

- + Any data domain plugs in via a connector
- + Whoever owns a domain brings it in – everyone benefits
- + Build the infrastructure once; law is just where we start

THE GOAL – TRENDING TOPICS

- + Surface the *pulse* of Switzerland's open data – emerging signals, not forecasts
- + **Pulse already runs:** 8.7M documents tokenized into a **49.6-billion-row frequency time-series (ClickHouse, ~2.3 TB)** + 8.1M embeddings – live trending today, you build on a working base
- + Today: law · next: politics, statistics, any domain – evidence-based & multilingual

LOOKING FOR

data-viz (D3) · NLP / topic-modeling · civic-tech curiosity