NicheExplorer

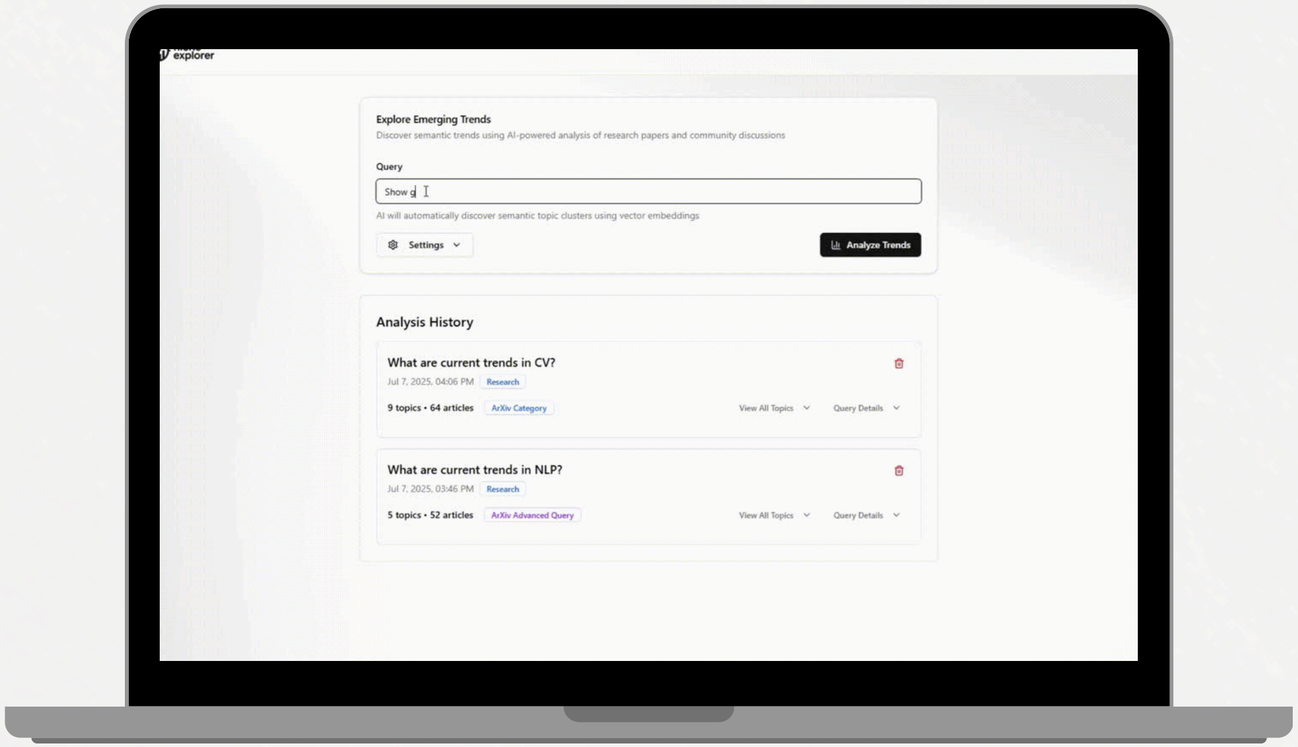
Niche Explorer - Janik Jehkul, Manuel Wilhelm

NicheExplorer

Problem: Great ideas are buried under a mountain of data

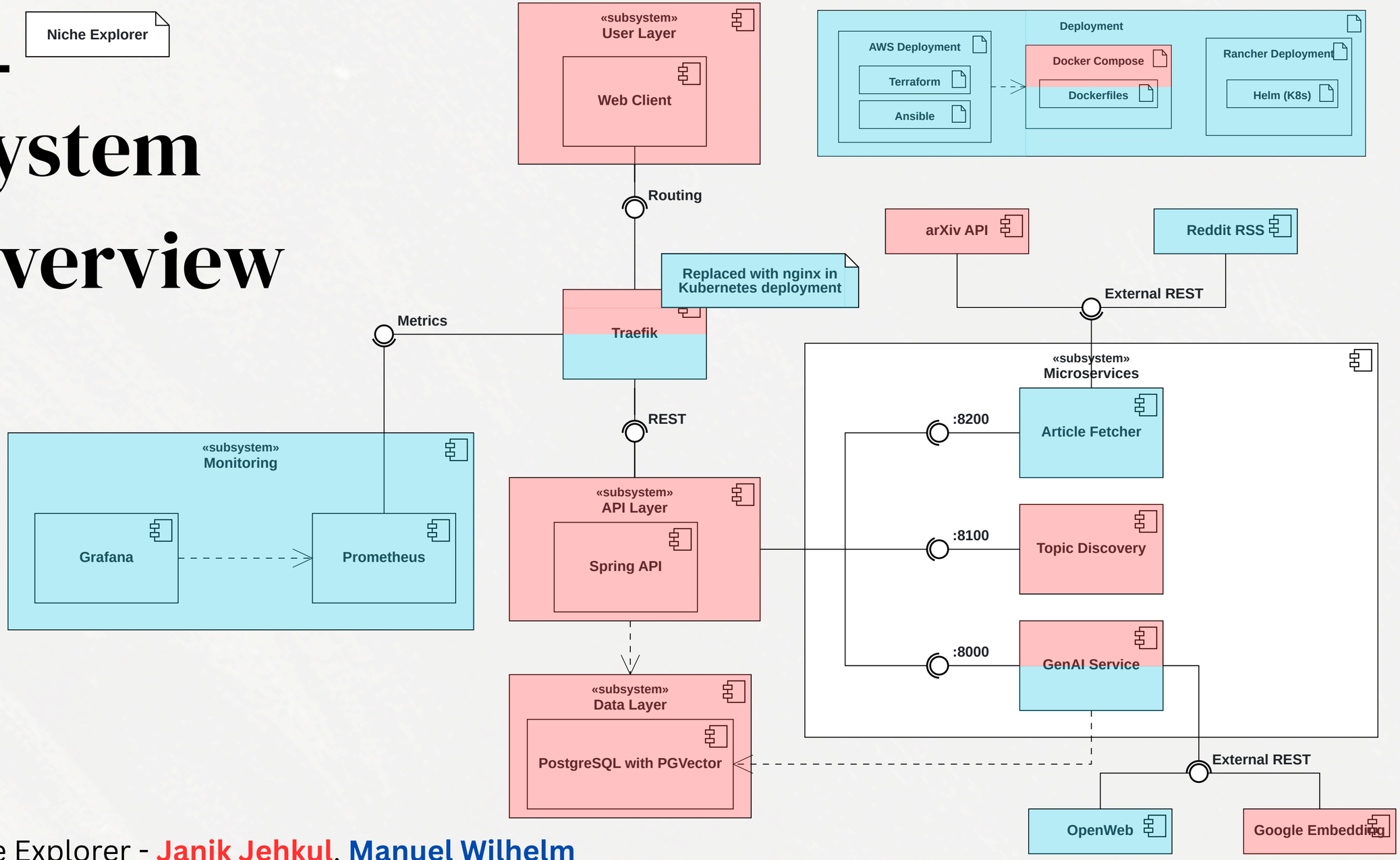
NicheExplorer:

- **Aggregates documents** from arXiv and reddit
- **Identifies underlying trends** using topic modeling
- **Presents the discovered topics** in an interactive list, allowing users to explore the topics



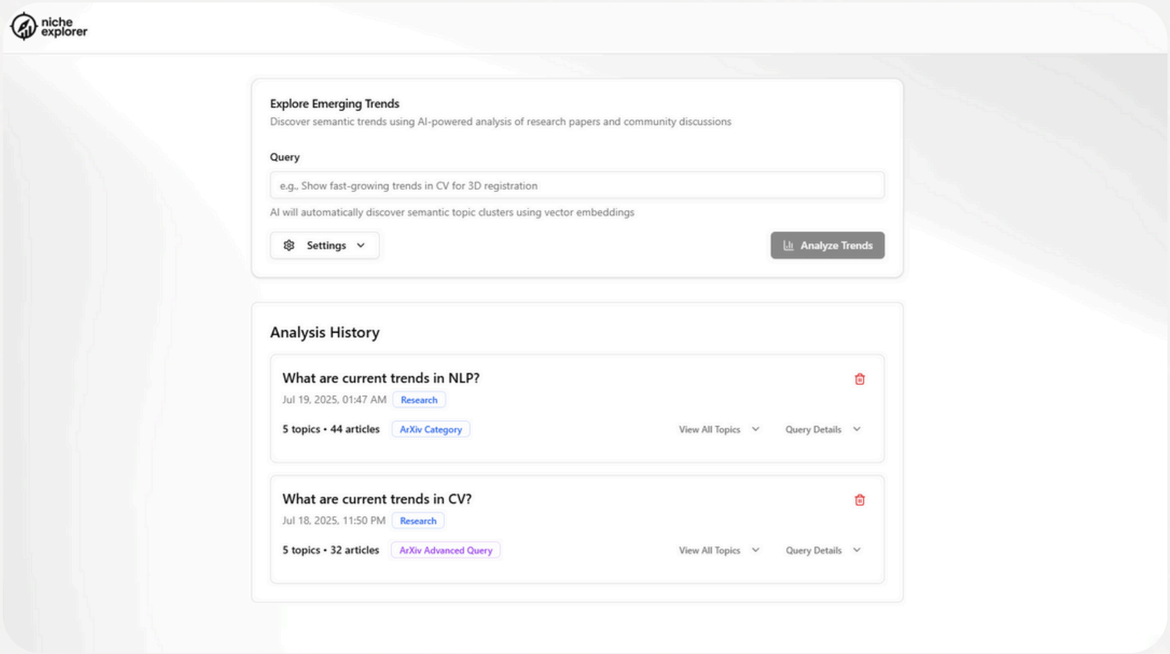
Niche Explorer

System Overview



Client - Frontend

- **Core stack:** React, Vite & TypeScript
- **UI Framework:** Tailwind CSS paired with shadcn components
- **Type-Safe API Contract:** Backend communication is handled by a strictly-typed API client auto-generated OpenAPI specification



Java

- **Core stack:** Spring Boot & Java, built with Gradle
- **Layered Architecture**:
  - **Controller**: validate & route HTTP requests/responses
  - **Services**:
    - **AnalysisService**: DB persistence (PostgreSQL + PGVector)
    - **AnalysisOrchestrationService**: microservice orchestration
  - **Config:** centralize API clients
- **Type-Safe API Contract:** Clients auto-generated from OpenAPI spec

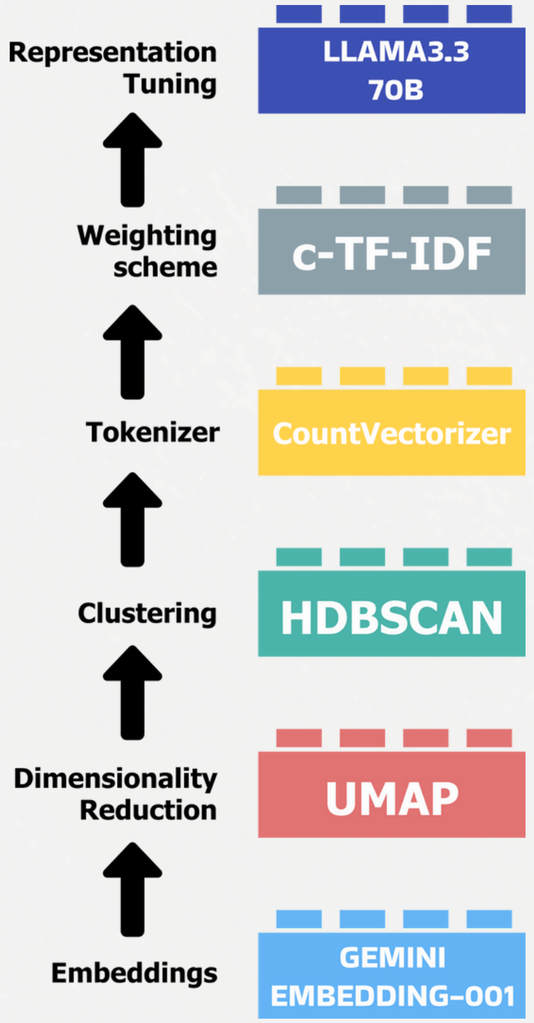
Python - Article Fetcher & Gen AI

- **Article Fetcher:** Python, FastAPI
  - Fetching ArXiv articles from the ArXiv API
  - Fetching Reddit posts from the Reddit RSS feed

- **Generative AI**:
  - **Text Generation**: Turns a prompt into structure text using OpenWebUI or Gemini
  - **Embedding**: Generates and fetches Gemini embeddings
  - **Classfication**: Predicts arXiv or reddit category & generates query

Python - Topic Discovery

- **Core stack:** Python, FastAPI
- **ML-Pipeline**:
  - Get embeddings from PGVector
  - Cluster articles with BERTopic
  - Sub-clustering when topic is large (>10 docs)
  - Label/summarize each topic
- **Output**: Ranked list of topics with title, description relevance



Based on: maartengr.github.io/BERTopic

Monitoring

Prometheus (Metrics, Alerts)

- **Core metrics:**
  - Service Availability
  - Number of Calls to External Providers
- Core **Alerts**:
  - Service Unavailable
  - Calls Approaching Provider Rate Limit
  - Calls Exceeding Provider Rate Limit

Grafana (Visualization)

Deployment

Local - Docker Compose	Docker Compose Local Docker Compose Overwrite
Rancher - K8s Deployment	Helm: Monitoring Deplyoment Application Deployment
Cloud - AWS Deployment	(Terraform) Ansible Docker Compose AWS Docker Compose Overwrite

Dropped out Teammember

- **When:** Midway through the project
- **Contribution:** No real contribution to the project
  - Lead to more work (major refactorings)
- **Missing parts of the project due to that reason**:
  - Query Responses in frontend will just be empty on errors / when finding nothing
  - Smoother project startup (2-5 minutes wait time before everything running)
  - More test (such as client tests)
  - Credentials management (Keycloak)

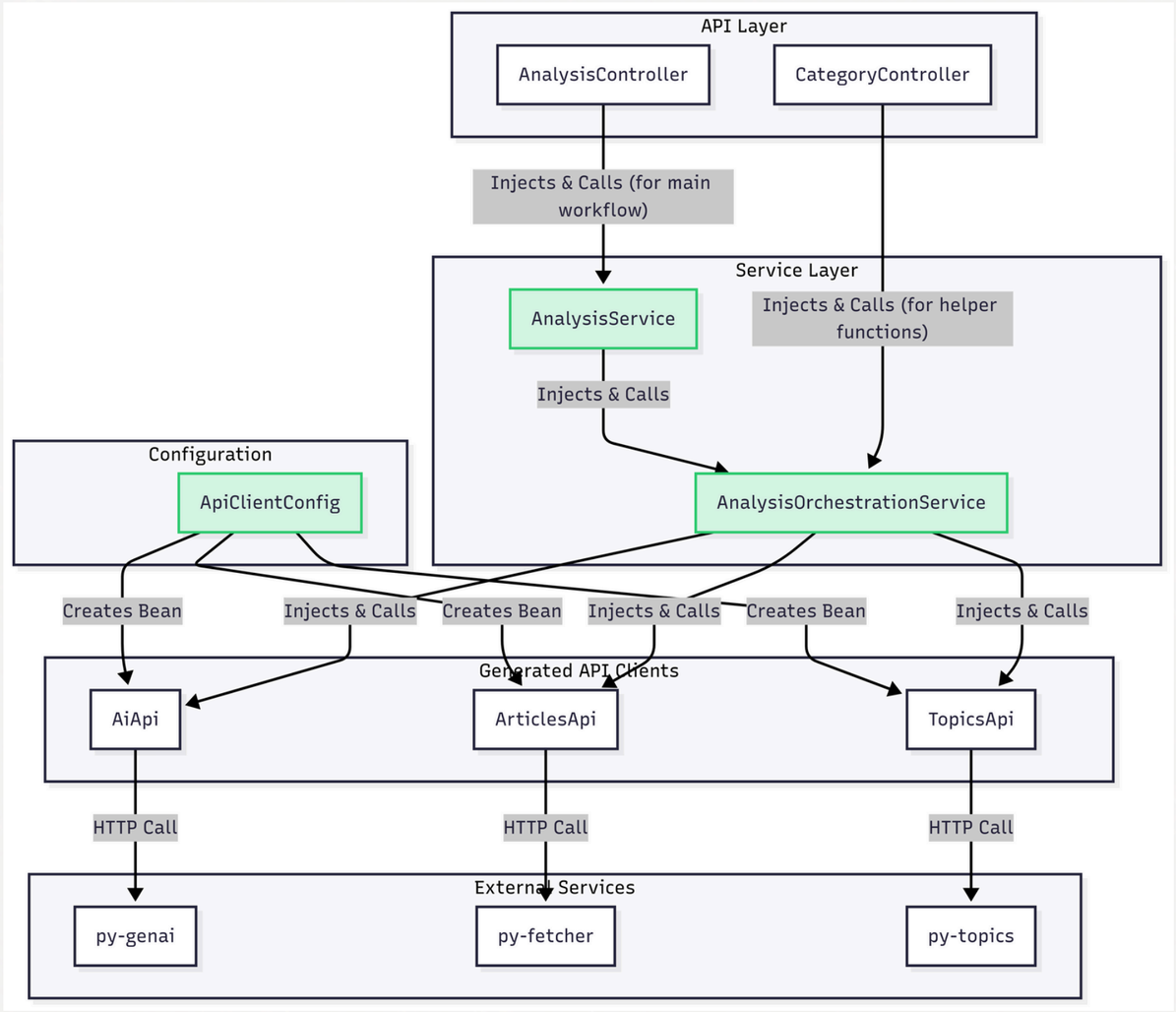
niche explorer

THANK YOU FOR LISTENING

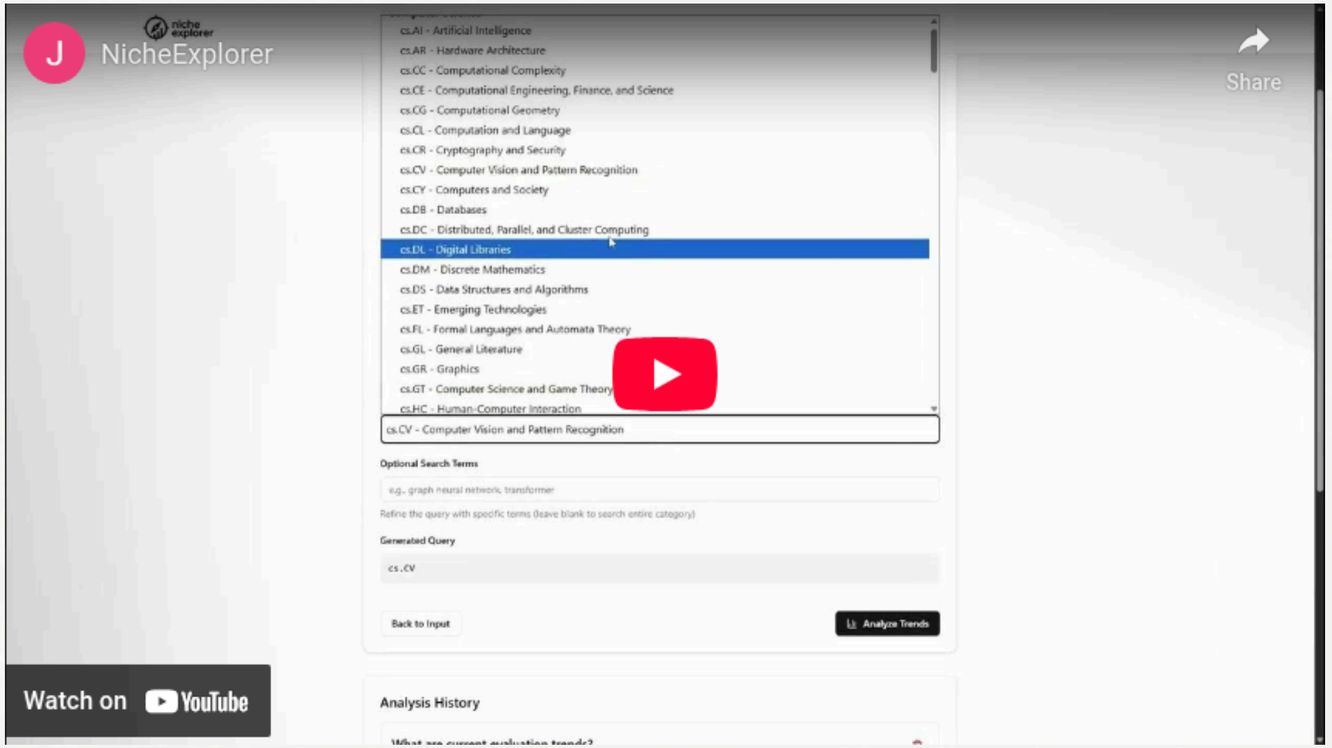
QUESTIONS?

Niche Explorer - Janik Jehkul, Manuel Wilhelm

Appendix - Java



Appendix - Demo



https://www.youtube.com/watch?v=JZuyDbpnB-A