SDMX REINVENTED WITH AI

From standards to insights: faster discovery, clearer visualization, smarter use

CONTEXT

Open-source IstatData platform is based on the international **SDMX standard and Istat Toolkit (M&DM and Data Browser)** accessed by over one million users annually.

~3.000	2 BILLION	19
Data tables	Aggregated records	Data Domains

User interaction is enhanced through an **Artificial Intelligent data search system** and **AI-powered table customization.**

GOAL

Enhance the use for those who do not have prior knowledge of the data, **simplify access**, **reduce contacts** to the contact center

RESULT

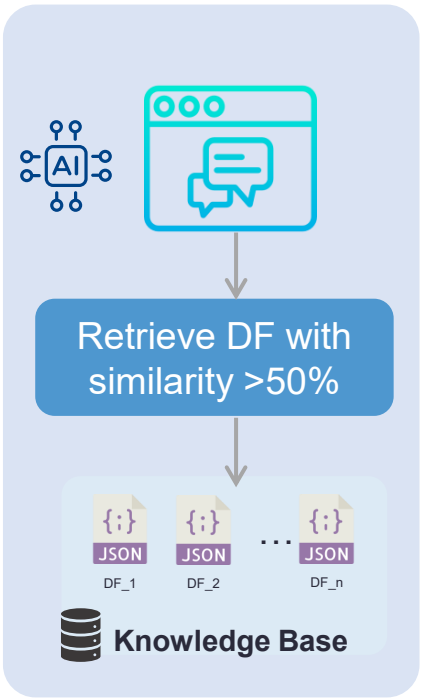
A **chat** based on **OpenAI** integrated into the portal that, with free requests in **natural language** and terms even non-existent in the database, allows access to data with **AI Visualization assistant.**

SEMANTIC SEARCH

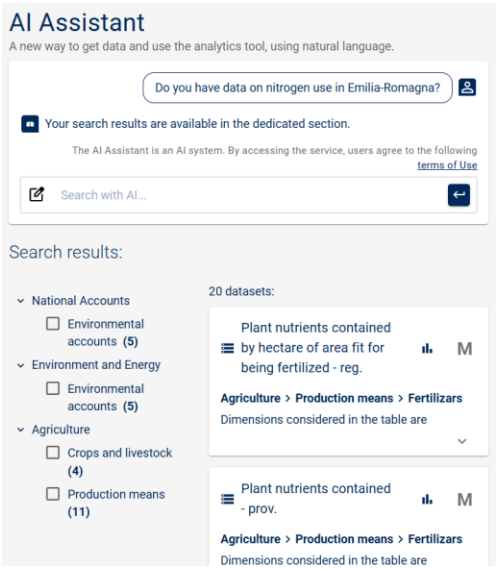
Understands natural language and returns the most relevant SDMX dataflows above a similarity threshold.

User Request

«Do you have data on nitrogen use in Emilia-Romagna?»



Subset of DFs ordered by relevance

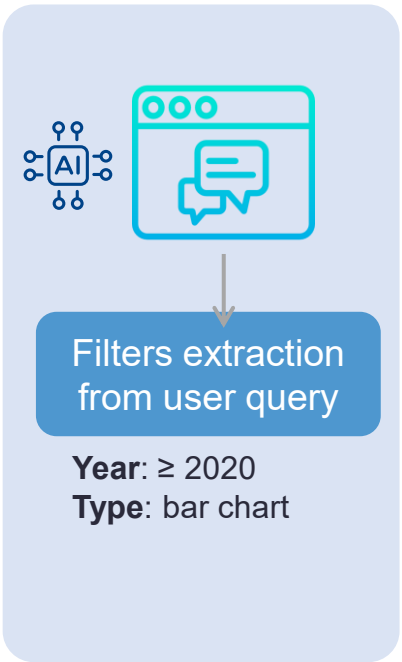


FROM SEARCH TO CUSTOM VIEWS

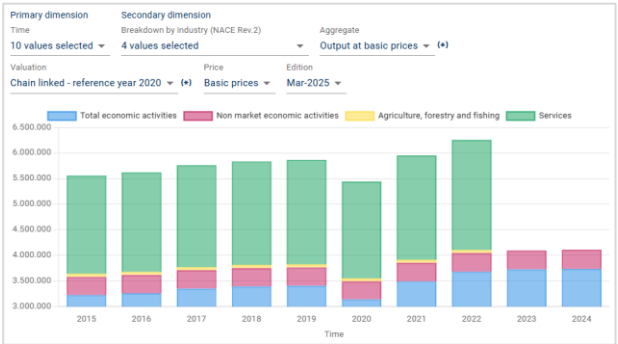
Use natural language to filter, pivot, and transform SDMX datasets into tailored tables, charts, or maps—with instant previews and guided insights

User Request

«Show me only the data from 2020 onwards, and display a bar chart»

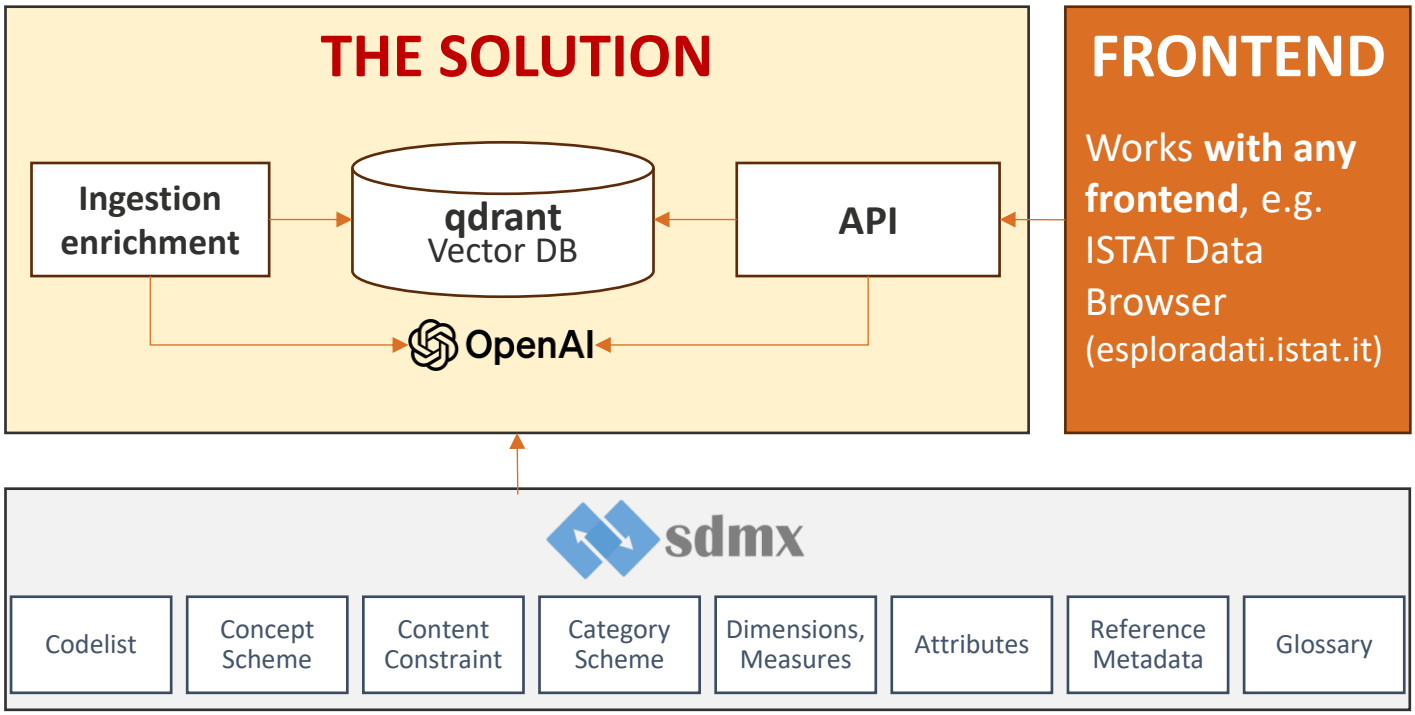


Time	2015	2016	2017	2018	2019	202
Breakdown by industry (NACE Rev.2)						
Non market economic activities	227.507,4	230.610	232.926,1	236.701,2	238.561,8	23
Agriculture, forestry and fishing	33.124,6	31.525,4	33.030,8	33.022,6	32.810	
Crop and animal production, hunting and related service activities	30.191,2	28.520,6	30.070,8	29.988,9	29.840,3	
Forestry and logging	1.915,1	2.000,4	2.041,1	2.103,9	2.159	
Fishing and	1.018,2	1.004,4	918,9	929,9	810,8	



HOW TO MAKE IT POSSIBLE?

An end-to-end architecture combining SDMX standards, semantic AI search, and customizable data visualization.



FEATURES

- **SDMX 2.1 REST/XML** support (3.0 on roadmap)
- **Resilient call handling**: cache, retry, rate limiting, circuit breaking
- **Semantic modeling** for complex domain enrichment
- **Linked and merged** datasets with **advanced constraints**
- **Hybrid search** (sparse + dense vectors) for top results
- Exact **glossary definitions** with similarity threshold
- Natural language switch: **chat, tables, charts, maps**
- Intent classification to **block unsafe or out-of-scope** queries
- **Delta updates**, parallelism, resilience, and monitoring
- **Deploy on** Docker, Kubernetes, OpenShift, IaaS, or bare metal
- Dynamic **configuration**: scheduling, re-indexing, threshold tuning
- **Multilingual accessible** UI with **progressive** chatbot responses
- Large-dimension handling with **reduction and fuzzy matching**
- **AI Act compliance** with transparent disclaimers

Solution Benefits and KPIs

Benefits are validated through objective performance indicators. Results highlight enhanced usability, reliability, and adoption

KPIS

~86%	~74%	~2.800
Search Precision @10	Search Precision @5	Weekly User Query

END-TO-END ADVANTAGES

For End Users

- Intuitive environment that hides SDMX complexity
- AI-powered assistant for natural language search and exploration
- Easy customization of tables, charts, and maps using natural language
- Access to glossary and guided explanations for full clarity

For Service Providers

- Standards-compliant, future-proof architecture (cloud-native or on-premise)
- Embraces modern data engineering and architecture practices
- Scalable, resilient, and fully monitored operations
- Faster onboarding and broader adoption across user communities

deda.next

This milestone is only the beginning of a broader journey