PaNFinder @ ICAT F2F

Nadine Fischer

ESRF

PANFINDER: THE F IN FAIR

F.A.I.R

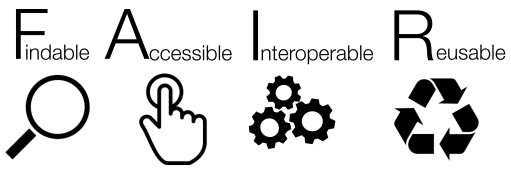
- ✓ Findable
- ✓ Accessible
- ✓ Interoperable
- ✓ Reusable

Without the F, your (meta) data is just AIR.

Finding the relevant data is the first step of (re)using data.

Powered by EU

Several EU initiatives to promote F.A.I.R principles, a corner stone to enable open science.



PANFINDER: HISTORY AND BACKGROUND

Predecessor - PaNOSC Data Portal

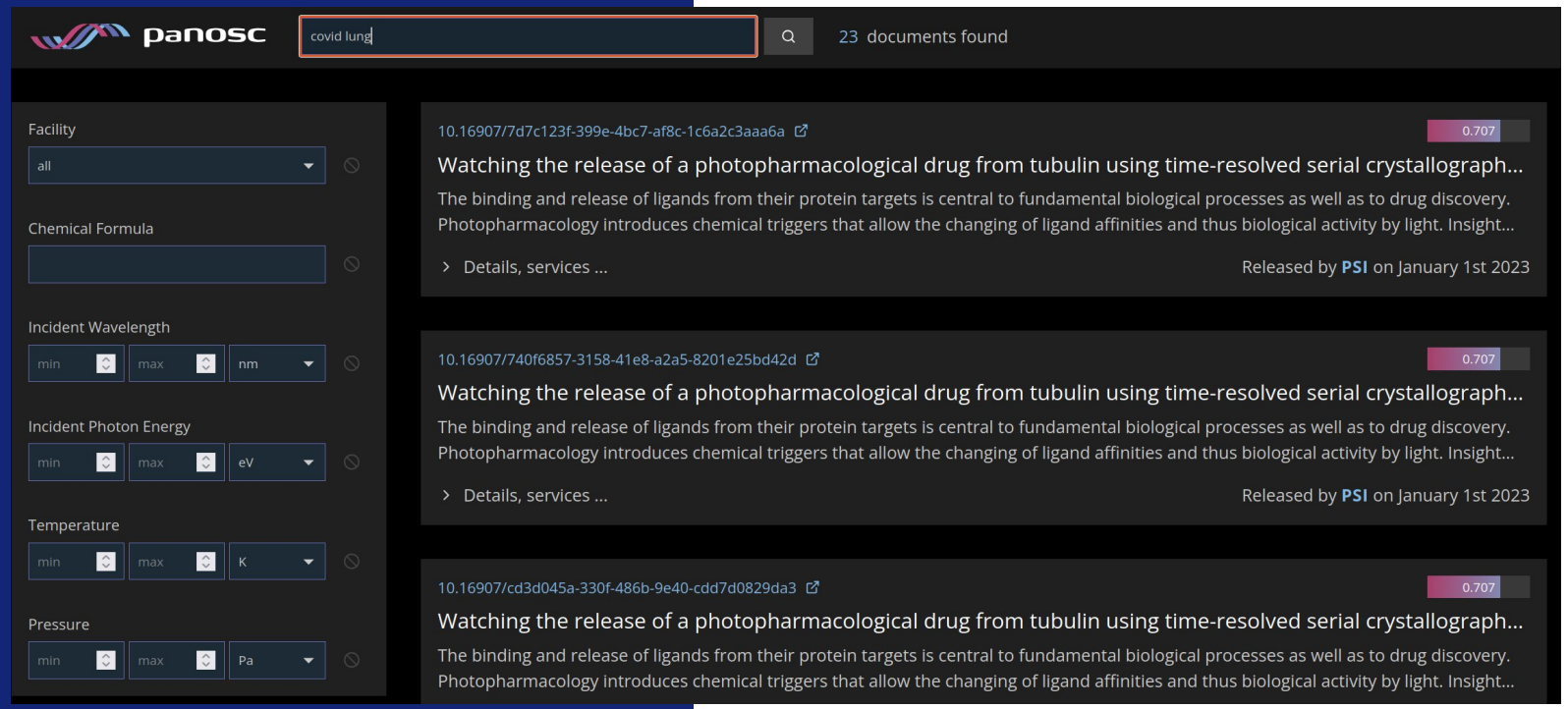
Federated search across multiple European neutron & photon facilities

- European Synchrotron Radiation Facility
- European Spallation Source
- Institut Laue Langevin
- MAX IV
- Paul Scherrer Institut
- Central European Research Infrastructure Consortium
- European XFEL
- ALBA Synchrotron
- Helmholtz-Zentrum Dresden-Rossendorf

Challenges

- **ExPaNDS project ended**
- **Unreliable connection between federated and local search APIs**
- **Meta data definition differs between facilities**
- **Non-intuitive user input**
- **Unresolved issues, e.g. filters**

✓ What is next?



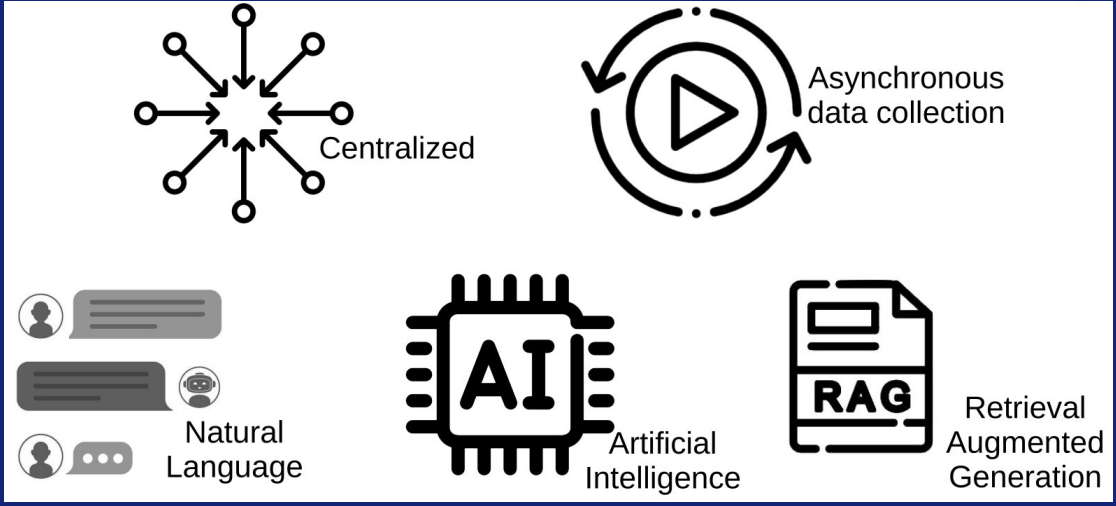
PANFINDER: THE BIRTH OF PANFINDER

Funding

- ✓ EOSC OSCARS
- ✓ Consolidates past achievements of the science clusters
- ✓ 2 Years, 2025 / 2026

OSCARS
Open Science Clusters' Action for Research & Society

Requirements



Team

- ✓ Max Novelli and the SIMS Team
- ✓ Notable mention: Janos Babik
- ✓ DMSC - ESS

ess EUROPEAN SPALLATION SOURCE

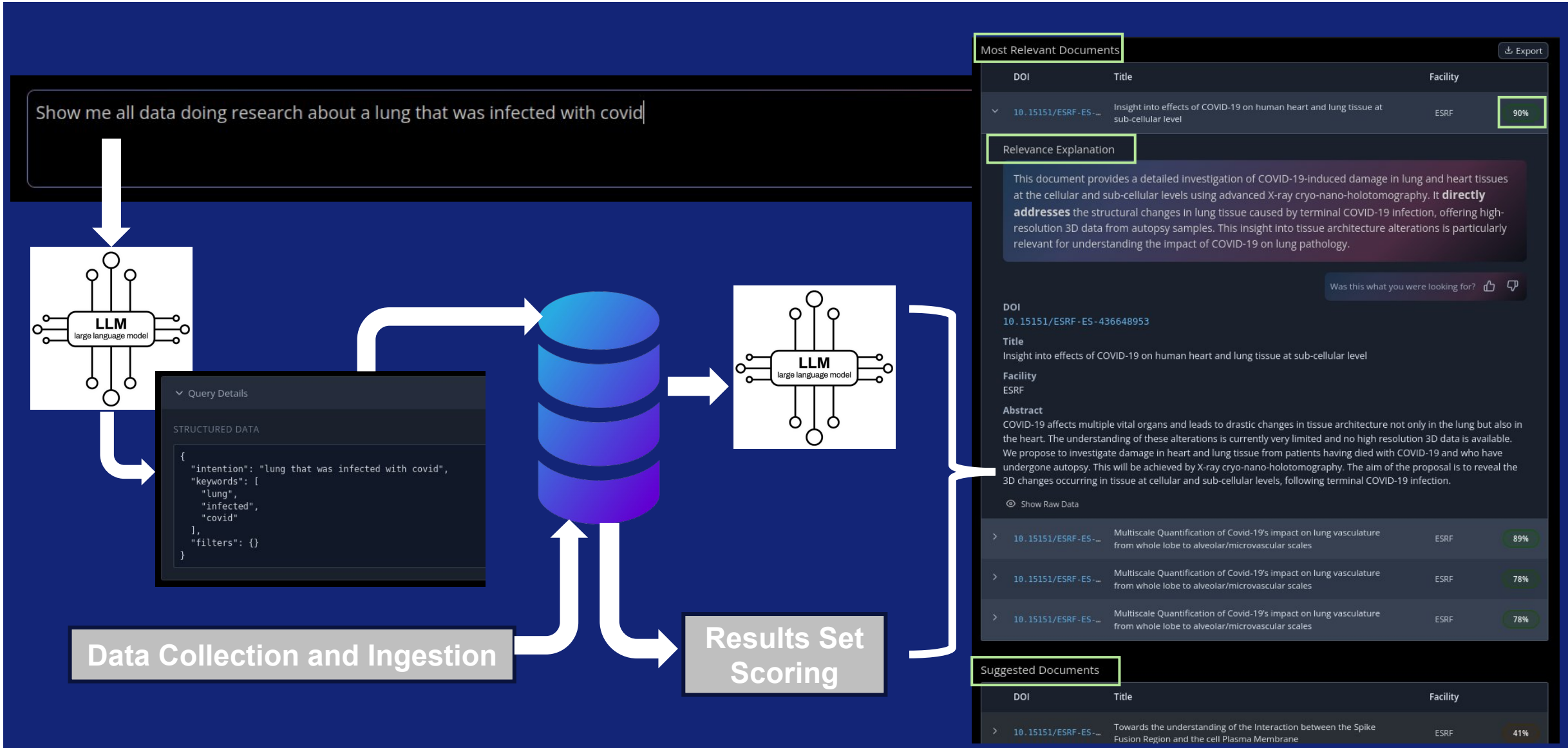
Idea

- Natural language query to find resources
  - URL, Datasets, Research Facilities
- ChatGPT-like interpretation and explanation of results

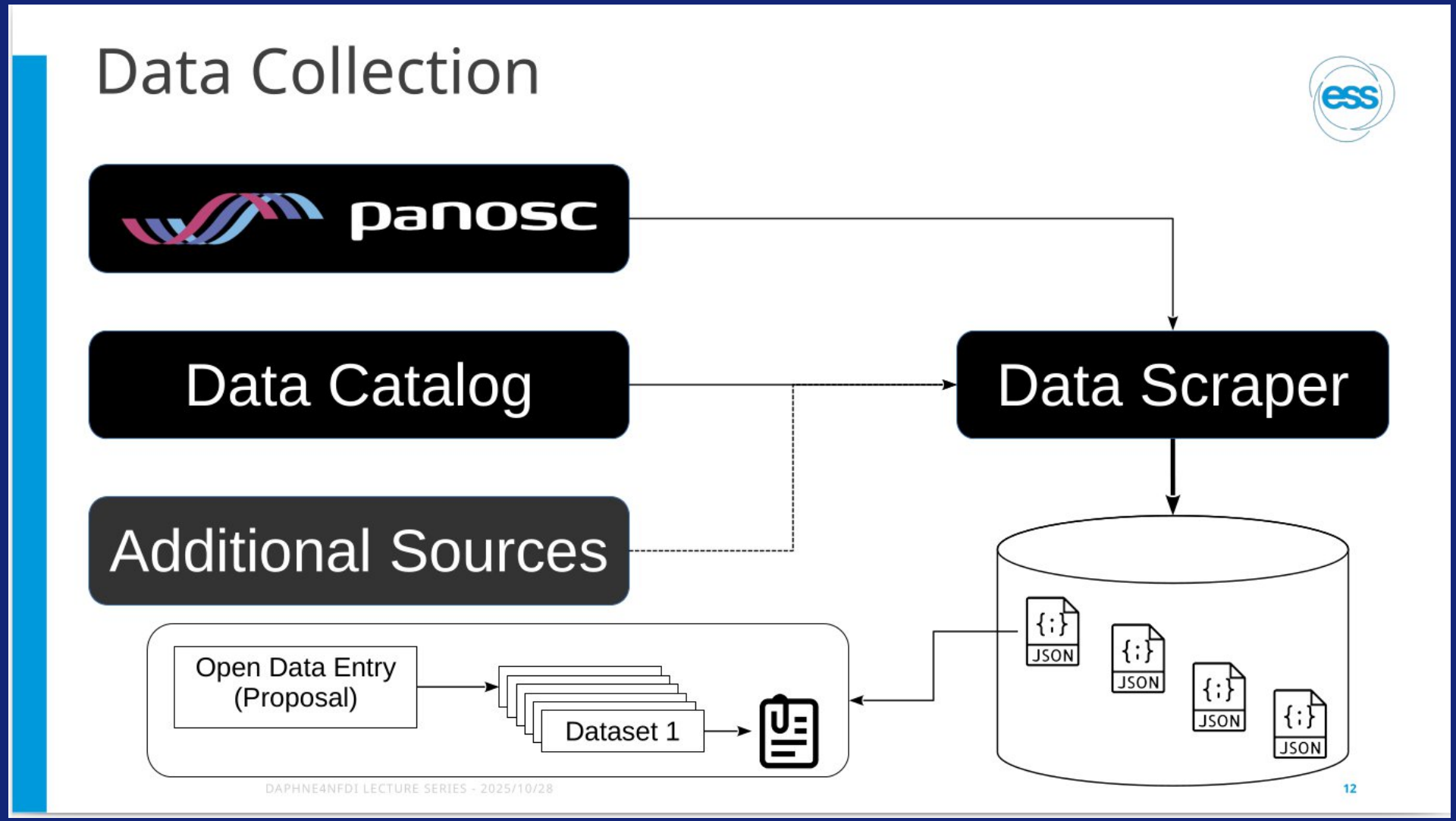
Current Status

- ✓ Beta Version
- ✓ https://pan-finder.panosc.eu
- ✓ https://pan-finder.panosc.ess.eu/pan-finder

PANFINDER: FUNCTIONALITY

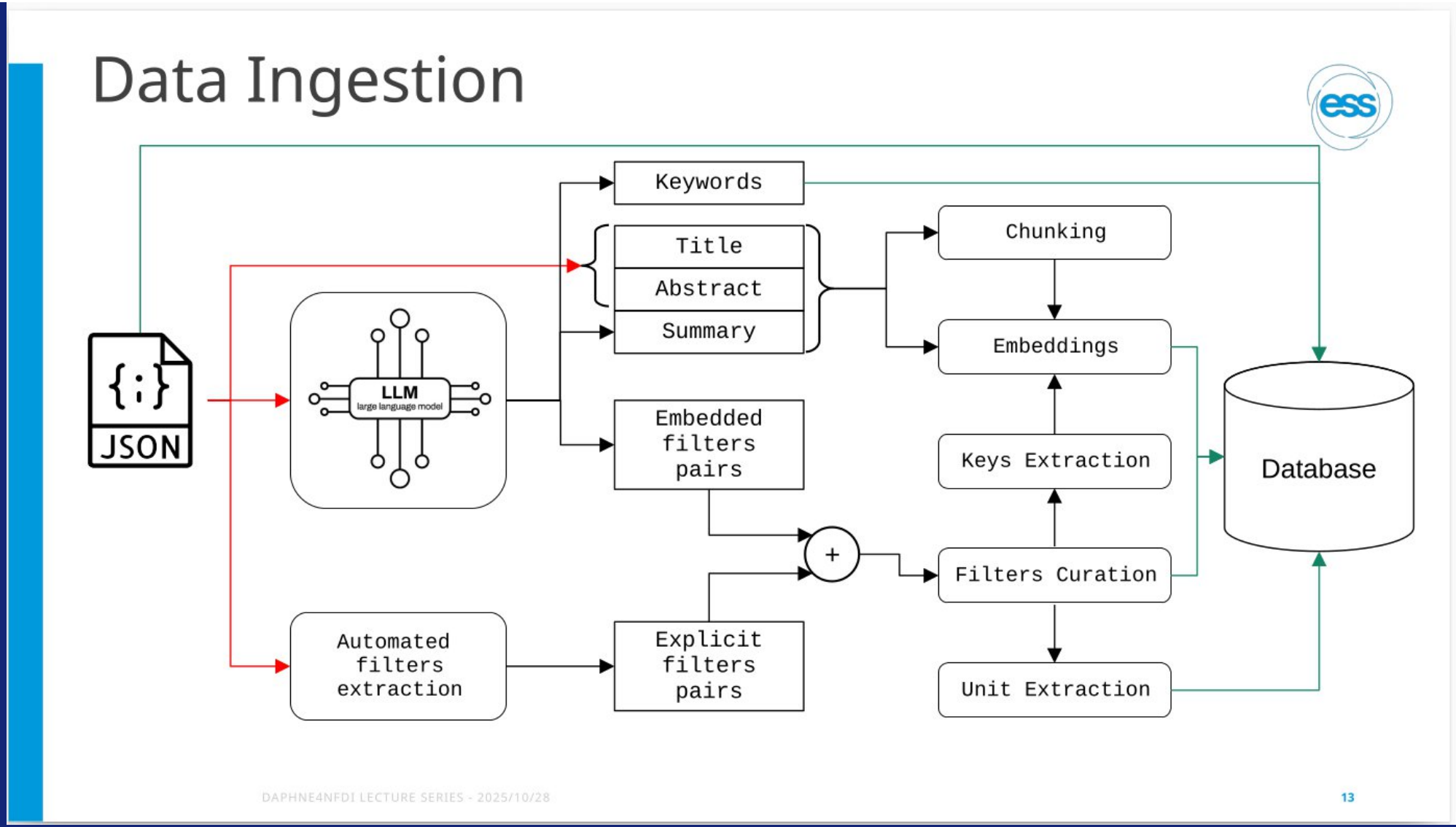


PANFINDER: FUNCTIONALITY



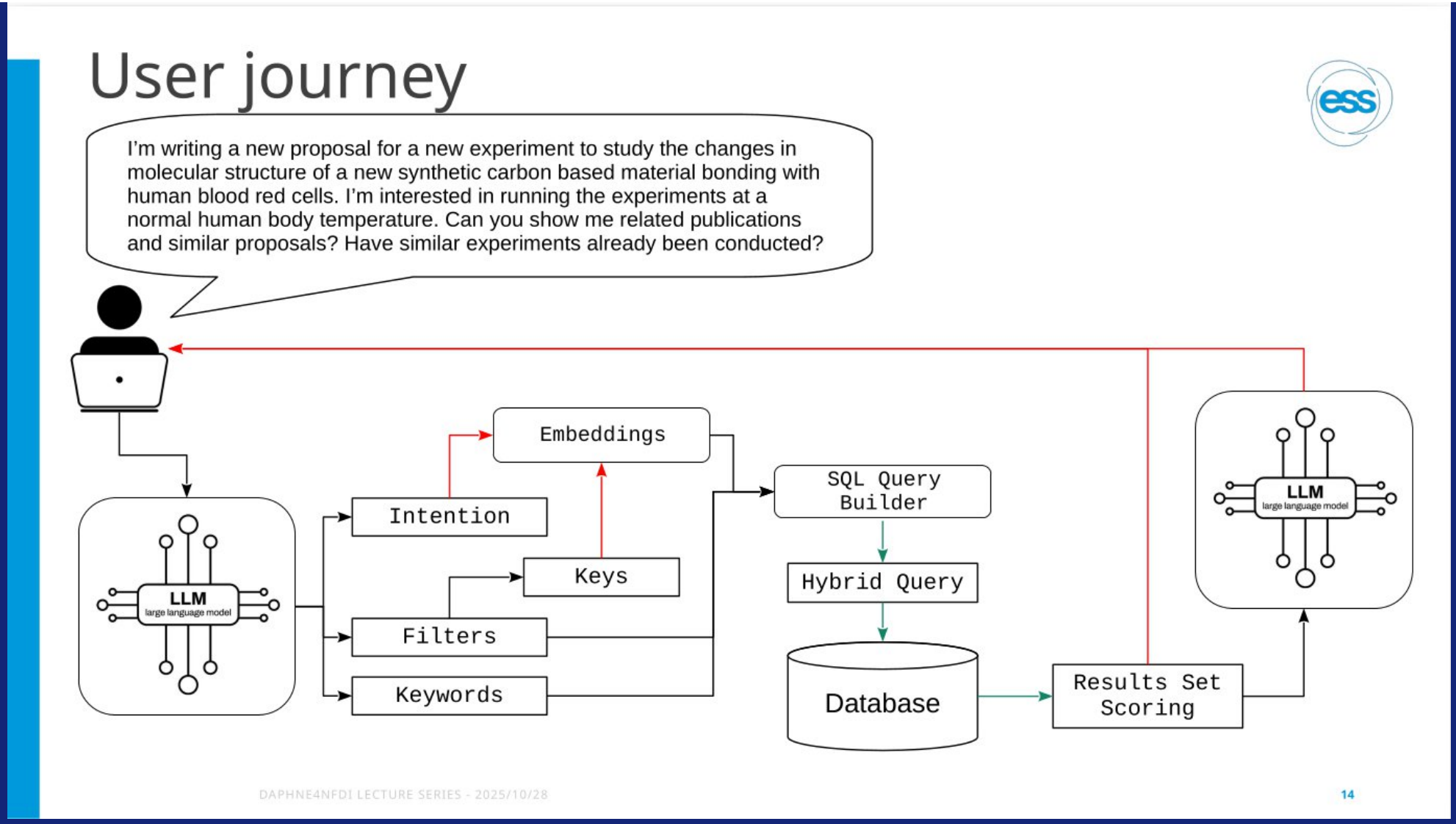
https://zenodo.org/records/17471792
Author : Max Novelli

PANFINDER: FUNCTIONALITY



https://zenodo.org/records/17471792
Author : Max Novelli

PANFINDER: FUNCTIONALITY



https://zenodo.org/records/17471792
Author : Max Novelli

PANFINDER: OUTLOOK

ESS

- Benchmarking
- More testing
  - Different models (size, cost, performance)
  - Prompt engineering
  - Model fine tuning
  - Refine natural language response
- Automating data collection

- User feedback:
  - 1-on-1 sessions, answers validation, and more
  - Volunteers needed!

- November 2026: Go public as part of deliverable

ESRF

- ✓ Two day knowledge transfer with Max and Janos (January 2026)
- Install PaNFinder locally and tailor to ESRF needs:
  - Manage authorization for embargoed data
  - Link documents to datasets OR samples
  - Utilize existing controlled dictionaries to standardize terminology
  - Implement live data ingestion via API
  - Test performance: currently thousands of documents, we would need hundreds of millions
- Additional ideas
  - ML for meta data with goals: decision making, quality or timing. Example: increase experiment efficiency and quality based on existing meta data
  - MCP (Model Context Protocol): Provides structured context to AI/ML models to improve analysis
  - LLM to generate the context for ML, MCP etc.

ACKNOWLEDGEMENTS

Some content and graphics from :
https://zenodo.org/records/17471792
Author : Max Novelli

Thanks to Andy Goetz, Marjolaine Bodin, Alejandro de Maria Antolinos and the whole DAU team for their continouus support !

QUESTIONS ?

CONTACT

nadine.fischer@esrf.fr

ESRF
The European Synchrotron

THANKS FOR YOUR ATTENTION