



EUROPEAN CENTRAL BANK

EUROSYSTEM

# Metadata integration as a key enabling technique for AI and ML

## IFC Workshop 2025

The views expressed are those of the authors and do not necessarily reflect those of the ECB.

20/02/2025



**Thomas Gottron, Daniel Suranyi**  
ECB Data Office

# Metadata supports and enables AI applications

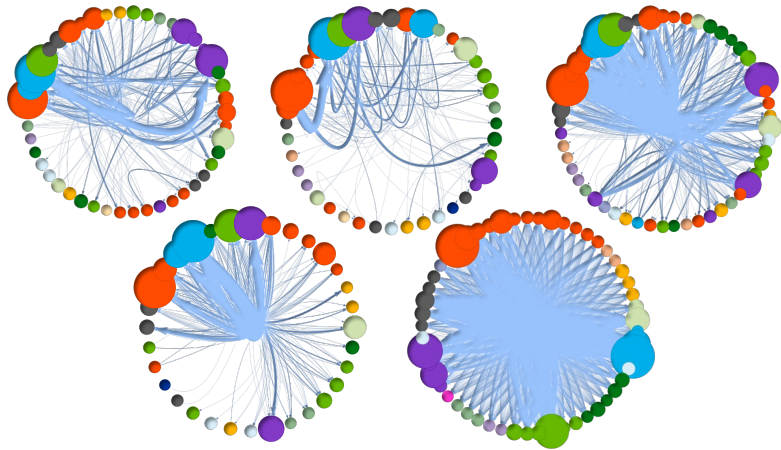


Metadata plays a crucial role for AI and ML:

- **Business metadata:** provision of context and methodological information
- **Semantic metadata:** integration, harmonisation and alignment of definitions
- **Structural metadata:** data models, classifications, mappings, transformations
- **Technical metadata:** insights on encodings, provenance, freshness, etc.

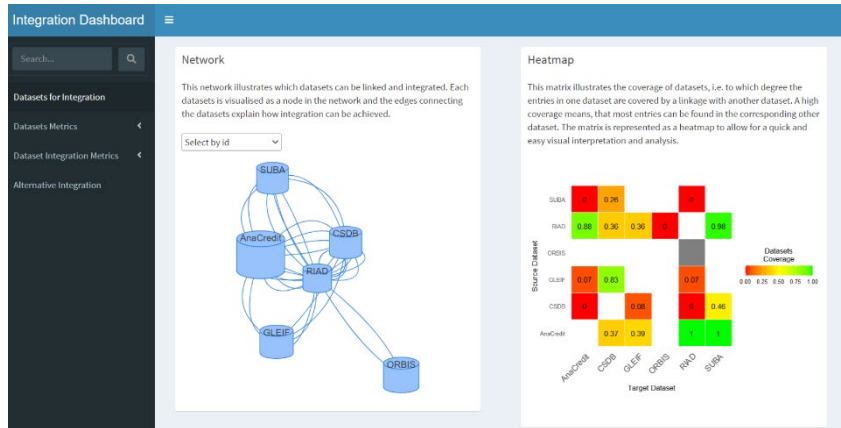
# Metadata driven AI use cases in central banking

## Analysis of bank exposures and systemic risks



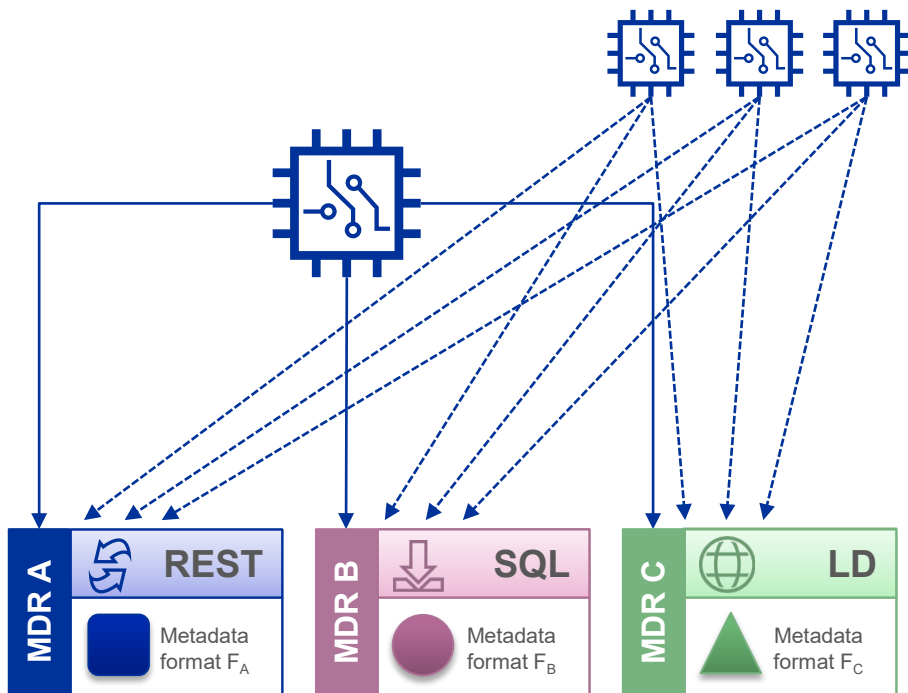
Aarab, I., & Gottron, T. (2024). Network Topology of the Euro Area Interbank Market. High-quality and Timely Statistics (pp. 3-19). Cham: Springer Nature Switzerland

## Quality assurance for integrated datasets



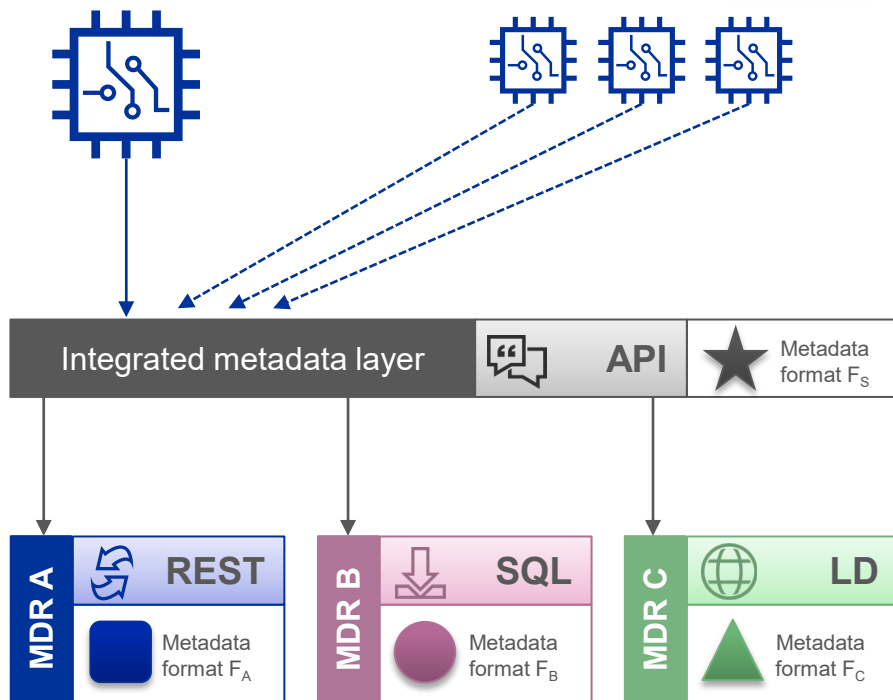
Gottron, T., Novello, A., Aarab, I., & Lauro, B. (2024). Assessing Fitness for Integration – a Metadata-driven Approach. 11th European Conference on Quality in Official Statistics (Q2024).

# Challenges in metadata integration



- Metadata silos even within organisations
- Lack of semantic integration and metadata governance
- Fragmented landscape of metadata standards
- Lack of standardised access to metadata repositories

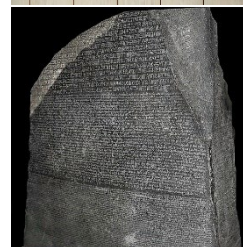
# Idea and benefits of an integrated metadata layer



Single metadata access point  
for all AI and ML applications



Standardised interface  
(API & format)



„Rosetta stone“ for  
translating between and  
connecting to heterogeneous  
metadata sources

# Key requirements for an integrated metadata layer

## Discovery, access & retrieval

- Standardised information model
- Modern standardised API for searching and retrieving metadata

## Support for different types of metadata

- Business metadata
- Semantic metadata
- Quality metadata
- Structural metadata, incl. master metadata

## Data integration and transformation

- Capability to describe granular, aggregated and master data
- Mappings between metadata definitions
- Support for data transformation

## Based on open standards

- Integrate with a wide ecosystem of solutions
- Extensibility and open to integration with external sources

# SDMX as backbone for an integrated metadata layer


*“It is felt that a clear, standard formatting of data and metadata for the purposes of exchange and dissemination can also facilitate internal processing by organizations and users, but this is not the focus of the specification”*

[SDMX Standards: Section 1](#)


Framework for SDMX technical standards, Version 3.0

# Assessment of SDMX and its features


## Discovery, access & retrieval

- 
- + SDMX registry
  - + Core standard information model
  - Fully fledged data catalogue will be more suitable for discovery


## Support for different types of metadata

- 
- + Support for reference data, structural information, description of data flows
  - + SDMX Metadata Structure Definition and Metadata Set

## Data integration and transformation

- 
- + Cross domain concepts and code lists
  - + Structure maps, representation maps, item scheme maps
  - + Integration of VTL

## Based on open standards

- 
- + SDMX as open and extensible ISO standard
  - + Wide range of existing third-party and open-source solutions and tools



# Summary and Conclusions

Harmonised and integrated metadata is a key enabler for AI and ML application



An integrated metadata layer helps to overcome drawbacks in current metadata landscapes



SDMX is a suitable solution to serve as backbone technology for an integrated metadata layer

