

# European Registers – sharing best practices

---

The ESCB's Master  
Dataset on Entities (RIAD)



29 June 2022  
CMFB workshop on BRs

Peter Neudorfer ([peter.neudorfer@ecb.europa.eu](mailto:peter.neudorfer@ecb.europa.eu))

RIAD team  
ECB DG-S

# The ESCB's Master Dataset on Entities (RIAD)

- 
- The 'Register of Institutions and Affiliates Data' (RIAD) comprises **Master Data information** converging to almost **12 mio** financial and non-financial **companies** ('legal entities')
  - Originally set up to support **statistics production** (register) ESCB wide (e.g. basis for the ECB **official lists** of MFIs, IFs, FVCs, PSRIs, ICs etc.)
  - Supporting multiple ESCB **business areas** (plus the SSM) including for instance a "**single data flow**" between **NCAs and EBA**
  - Facing increasing demand to support **daily operations** of various ESCB processes
  - Next **major extension** with the **IReF** (ESCB integrated reporting framework)

# Outline

■ Stylised facts

Unique identification

Compounding

Relationships and groups

Confidentiality and data access

■ IReF challenges

■ Way forward

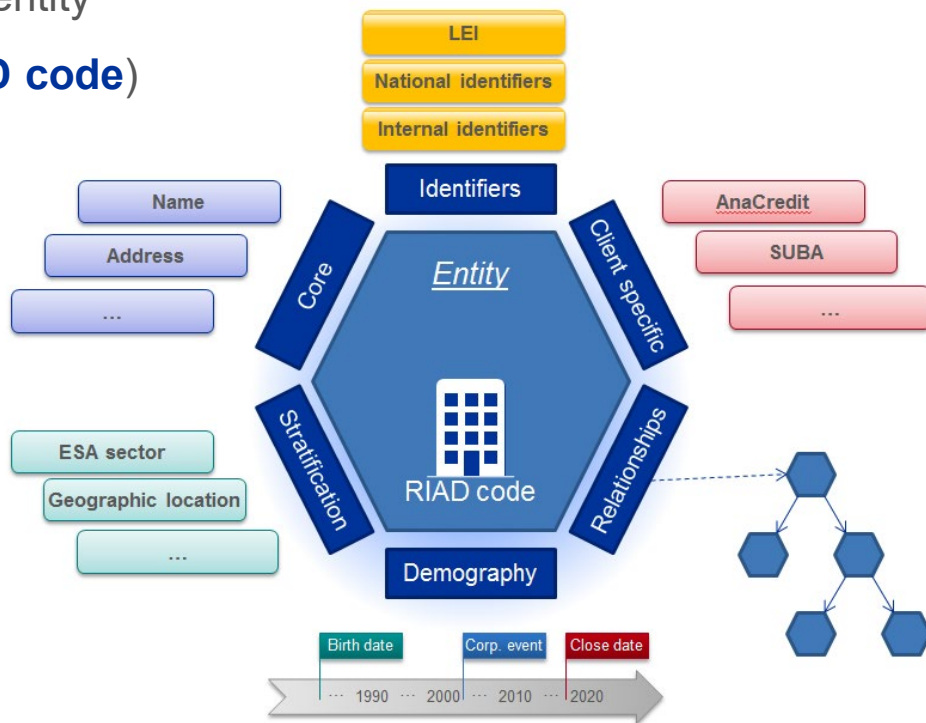
# Stylized facts

# A 'Master Data' set shared among NCBs/NCAs

- **RIAD clients *retrieve*** data from RIAD but also ***contribute*** data to RIAD, i.e. they
  - ensure the necessary sourcing of relevant master databased on legal acts (e.g. Guidelines ECB/2018/16 and ECB/2020/16 , Decisions etc.)
  - record new relevant entities, addition of necessary information (incl. on relationships)
  - perform the necessary DQM on client-specific data
- **RIAD Hub Network (RHN) & ECB RIAD Team** take care of
  - unique identification of counterparties
  - methodological soundness (e.g. via the 'Single Data Dictionary' (SDD) )
  - DQM on core (identification / stratification) data
  - system availability and data accessibility

# RIAD data model

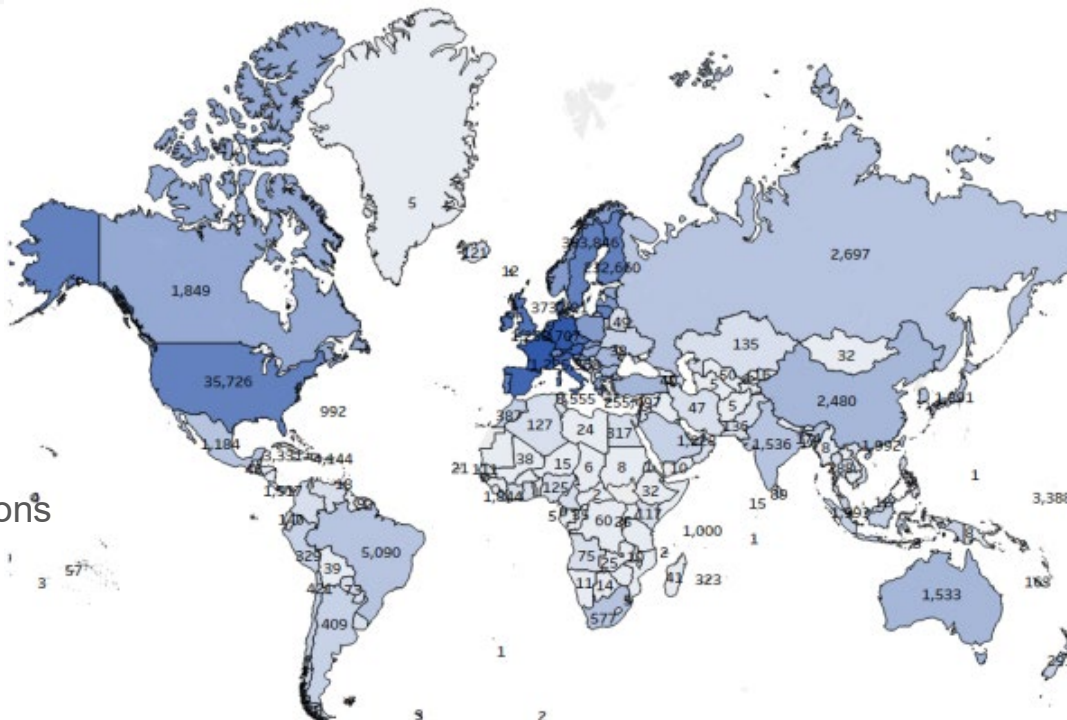
- Covers around **100** information items per entity
- Core ‘**identification**’ properties (incl. **RIAD code**)
- Allows matching of **193 identifier types**
- **Stratification** data
- **Client-specific** modules
- **Demography** and change events
- **Relationships** (subsidiaries etc. to built **groups**)
- “**Double Historisation**”
- **Daily updates**



## What you find in RIAD now

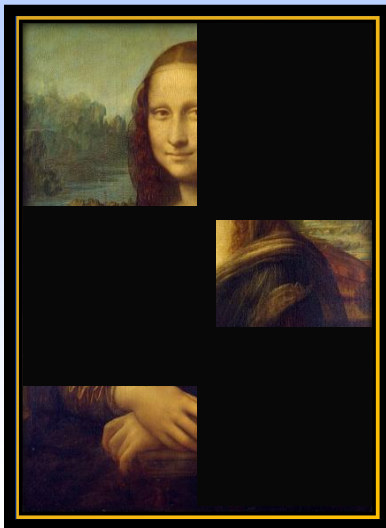
Includes around **11,5 million entities:**

- all EU banks and other financial institutions
- Non-financial entities (incl. SMEs) as demanded by AnaCredit
- 0,8% resident outside the EU



# Key data quality challenges

## Completeness



The information about an entity is complete

## Consistency



Data values are free of contradictions

## Uniqueness

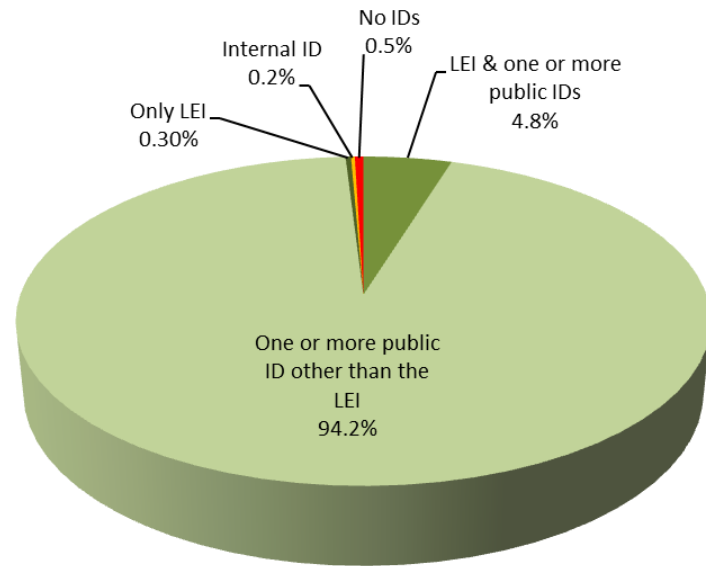


There are no duplicates of an entity

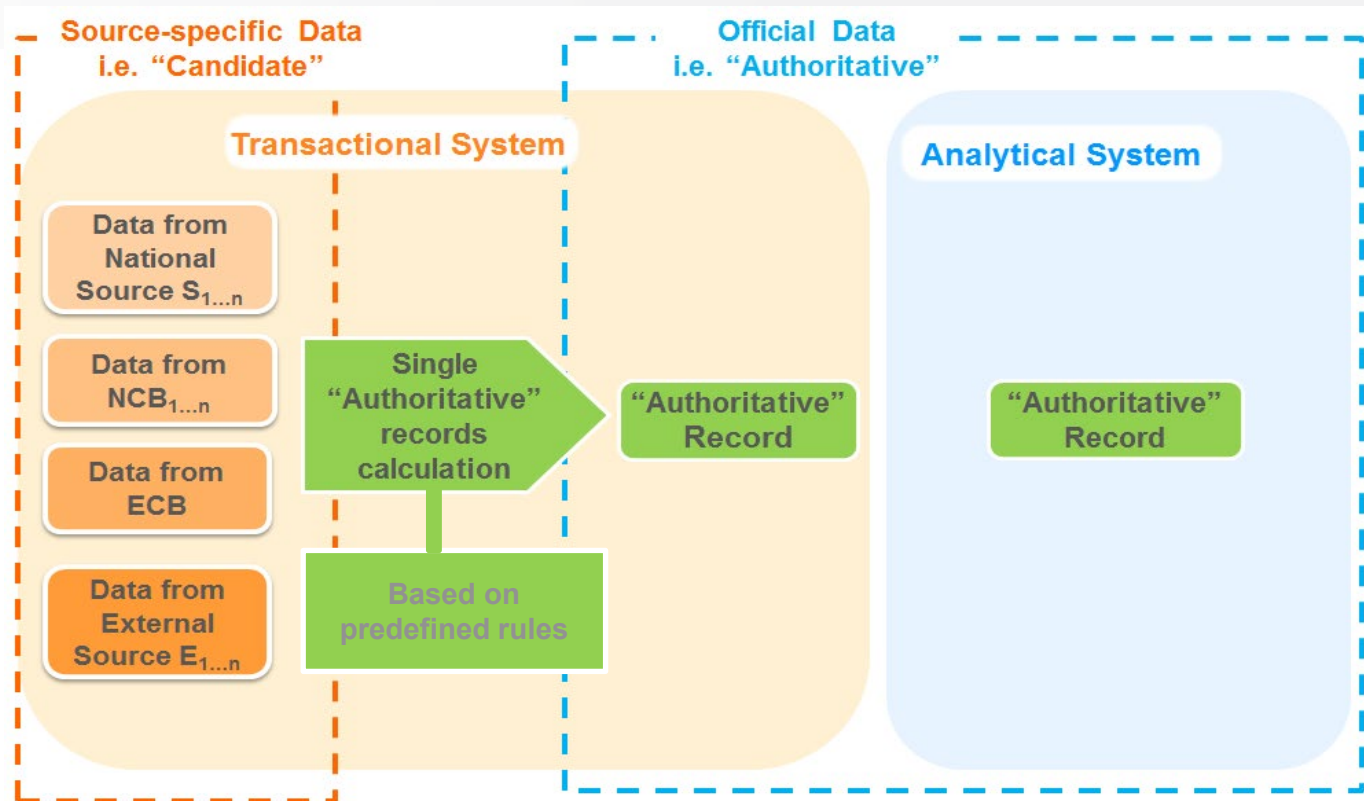


# Entity identifiers: the treasure of RIAD

- Entity identifiers play **pivotal role** in RIAD
  - ✓ at **least one ID** must be provided at registration
  - ✓ NCBs are requested to upload all available IDs - **more than one** - for each resident entity
- RIAD can store 193 **pre-defined** identifier types + (non-pre-defined) **'other' identifiers**
- Almost complete coverage → **99.3%** of entities have at least one **publicly available identifier**
- The desired **high connectivity** towards other datasets - a key analytical value – would in particular benefit from a wider assignment of LEI codes



# ‘Authoritative’ record building in a ‘multi source’ set-up



# From relationships to groups

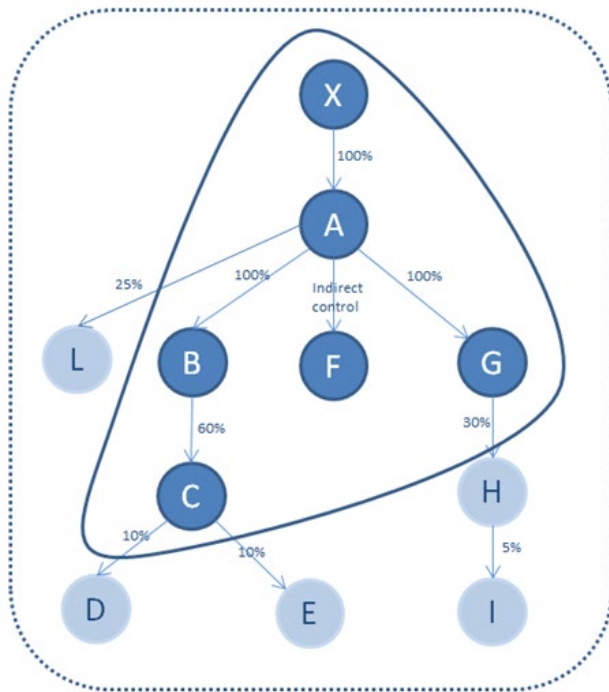
In RIAD **group data** are not exogenously loaded but **daily calculated** based on **relationships between entities** (e.g. *Is-owner-of*, *Is-branch-of*)

- **Advantages**

- Flexibility in modelling groups
- Different views possible, depending on users' need
- Full hierarchy of the group ('group structure')

- **Challenges**

- Information on relationships is hard to get
- Hard to disentangle mistakes from intended changes
- High demand for more complete group structures, e.g. including more NFCs
- Timeliness of changes



# Confidentiality and data access

- Each **value** in RIAD is associated to its own **confidentiality status**:
  - **F**: (potentially) free for publication
  - **R**: can be shared with reporting agents
  - **N**: cannot be disseminated outside the ESCB
  - **C**: confidential (ECB restricted) data
- In addition, specific segments of RIAD are only accessible to **selected user groups**
- The **access management** include special application **roles** and different levels of **confidentiality**
- Overall goal that (at least the) **core attributes** are flagged **F** or **R** → in particular, the IReF initiative strongly relies on the shareability of master data with reporting agents the for regulatory reporting

# Beyond the role of a “register for statistics compilation”

- Besides its role in **statistics compilation** and **data analytics**, RIAD as the authoritative source on entities in the ESCB is more and more key for several **operational processes**
- **RIAD** data is increasingly supporting key **ECB business processes** (CeBa & SSM)
- **Continues evolution** triggered by a broad set of client systems and users
- Compared to the role of a “register supporting statistics production” this comes with stronger demand for **timeliness** & **accuracy**.
- Already now RIAD provides daily updates in an overnight process. In the future an even **higher frequency (of provision) of updates** may be required – potentially intraday updates for systems relying on strict end-of day processing

# Challenges of the “Integrated Reporting Framework” (IReF)

# IReF demand for RIAD

- The **IReF** seeks to **integrate** the existing ESCB **statistical data requirements for banks** into a unique and standardised reporting framework applicable across the euro area.
- Covering current regulations on BSI and MIR statistics, SHS-S or AnaCredit (and other Guidelines).
- **Huge impact on RIAD** (processes, IT systems, etc.)
  - Need to make (some) data **available to reporting agents to support regular reporting**: challenges for IT system, ***data confidentiality***
  - Need to remove **legal and commercial obstacles** at national level hampering data sharing
  - **Data volume** / entity population may **grow** significantly (up to 20-30 million entities?))
  - High **data quality** required for properties in the scope of the IReF

→ Will materialise in a **new version RIAD 5.0**

# Way forward



# Focus of the next years

---

Expand role  
as ESCB's  
**Master  
Dataset** on  
Entities

Supporting key ESCB business processes with  
**daily** and (potential even) **intra-day data processing**

---

Deliver the functions required by IReF (**RIAD 5.0**),  
including expected data sharing with reporting agents  
for regulatory reporting

---

While addressing the above needs, investigate further  
synergies with **European Institutions** (EBA / SRB /  
EIOPA / ESMA / Eurostat)

---

# Thank you!