EUROPEAN CENTRAL BANK

EUROSYSTEM

Silvia Giacinti
Principal economist
DG-Statistics
Analytical Credit & Master Data Division

Single Data Dictionary (SDD) and Banks' Integrated reporting dictionary (BIRD)

Eurofiling conference 19 June 2019

- ESCB statistics to keep *supporting policy making* while *minimising the reporting burden for banks* via:

  – Engaging in a regular dialogue with the banking industry

  – Standardising and integrating existing frameworks for banks' reporting across *domains* and across *countries*

IReF

Integrated Reporting Framework

Collect the data only once, via an integrated reporting scheme

Focus on *ESCB statistical requirements*

BIRD

Banks' Integrated Reporting Dictionary

Support reporting agents to optimally organise the information available in their internal systems

ESCB strategy on data integration 2/2

Semantic integration

Semantic and reporting integration

BANKS

EUROPEAN AUTHORITIES

Operational systems

Transformations by banks

BIRD input layer

Transformations defined by banks and authorities

IReF collection layer (*)

Transformations by the ESCB

ESCB statistics

IReF

SDD

Transformations defined by banks and authorities

Supervisory/resolution reporting
EBA, SSM and SRM

BIRD (**)

Specific SDD framework

3

* **Using the ECB Single data Dictionary**

www.ecb.europa.eu ©

**See https://banks-integrated-reporting-dictionary.eu/

SMCube Methodology
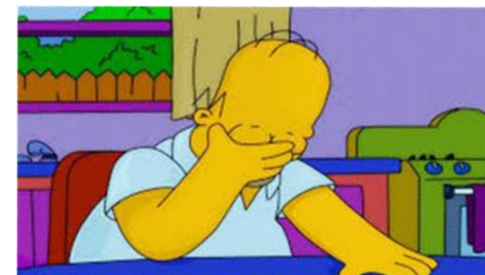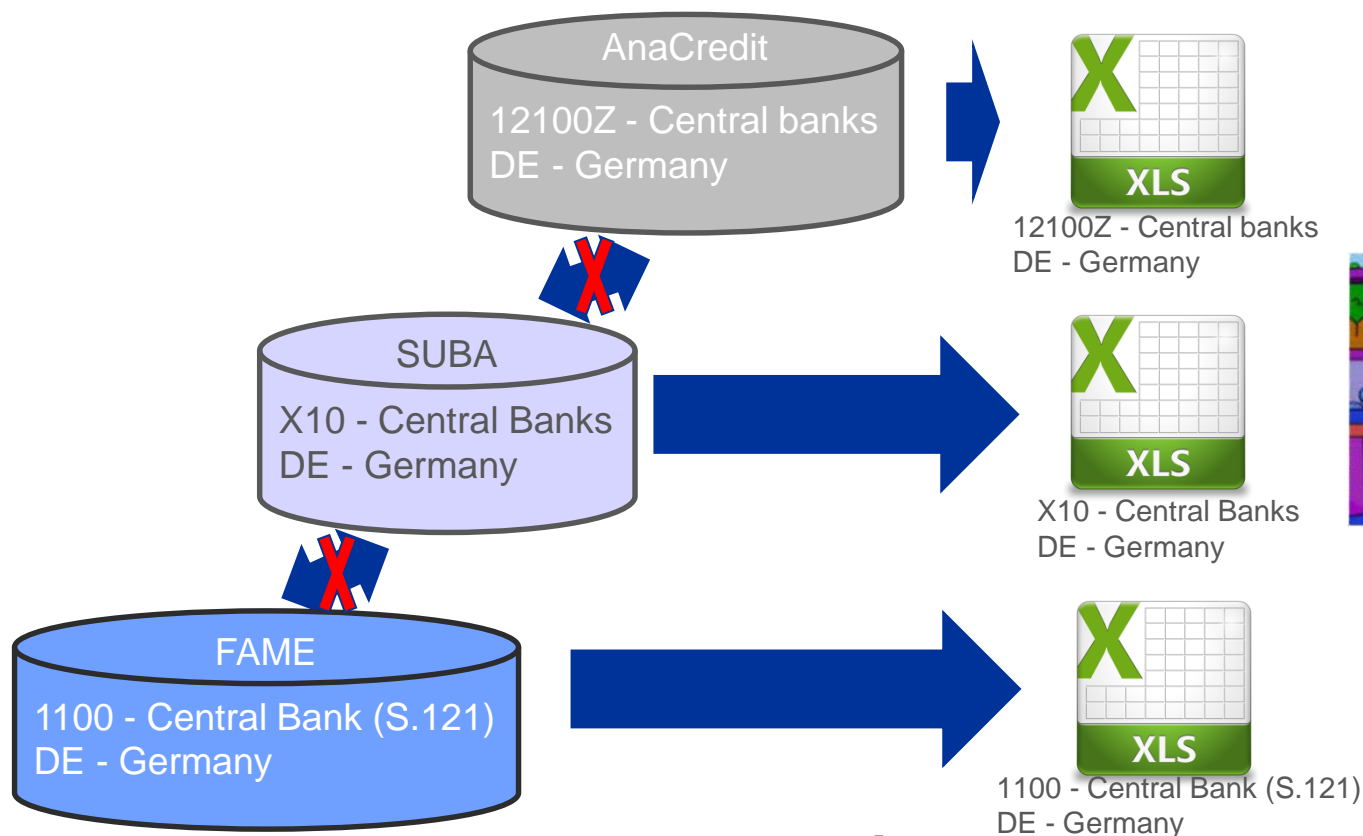
The BIRD and the SDD are based on SMCube:

- New methodology developed in ECB to define metadata and describe datasets.
- You can describe exactly:
    - <u>What are the fields (columns) of the dataset</u>
    - <u>What are the allowed values for each field</u>
    - <u>Mapping different metadata dictionaries like DPM and SDMX registry</u>
- Why a new methodology?
    - Because in ECB we need to bridge different metadata dictionaries (mainly SDMX and DPM) and combine dataset belonging to different metadata dictionaries
    - SDMX was not able to contain DPM metadata, and DPM was not able to contain SDMX metadata

New methodology

The lack of methodological integration solved with SMCube

Are definitions in FinRep, BSI and AnaCredit consistent ?

AnaCredit
12100Z - Central banks
DE - Germany

12100Z - Central banks
DE - Germany

SUBA
X10 - Central Banks
DE - Germany

X10 - Central Banks
DE - Germany

FAME
1100 - Central Bank (S.121)
DE - Germany

1100 - Central Bank (S.121)
DE - Germany

Need for semantic integration
Current situation *(Before conversion)*

SHS – Holdings of securities

REFERENCE_ PERIOD	IDENTIFIER_ TYPE	IDENTIFIER_ VALUE	HOLDING_ AMOUNT_TYPE	HOLDING_ VALUATION_ TYPE	ISSUER_DO MICILE_ COUNTRY	INSTRUMENT _ESA_ CLASS	HOLDING_ AMOUNT_ VALUE
Reference Period	Identifier Type	Identifier Value	Amount Type	Valuation Type	Issuer Country	Instrument Esa Class	Holding Amount
31/12/2016	ISIN	XX1234567	LE	M	DE	F_32	2000
31/12/2016	ISIN	XX9876543	LE	M	DE	F_511	900

CSDB – Securities issued

CORRECTIONDATE	EXTERNALCODE_ISIN	IDIRCOUNTRY	IDIRCLASSIFICATIONCODE_ESAI10	PUBLICATIONPRICE
Correction date	ISIN code	Issuer country	ESA 2010 instrument class.	AVG Publication price
31/01/2017	XX1234567	DE	F_32	2200
31/01/2017	XX9876543	DE	F_511	3000

Code (Memb..	INSTR_CLAS
F_31	Short-term debt securities
F_32	Long-term debt securities
F_511	Listed shares
F_512	Unlisted shares

KA	Other changes in the volume of assets
LE	Position (Stock)
NP	Number of purchases transactions
AC	At amortised cost
M	Market value

6

# Need for semantic integration
After semantic integration

SHS reference

DT_RFRNC	ISIN	CNTRY_ISSR	TYP_INSTRMNT	ORGNL_MTRTY	MRKT_VL
Reference date	ISIN code	Country of the issuer	Type of instrument	Original maturity	Market value
31/12/2016	XX1234567	DE	210 (Debt instrument)	2 (Short term)	2000
31/12/2016	XX9876543	DE	317 (Equity instrument)	0 (not applicable)	900

CSDB reference

DT_RFRNC	ISIN	CNTRY_ISSR	TYP_INSTRMNT	ORGNL_MTRTY	MRKT_VL
Reference date	ISIN code	Country of the issuer	Type of instrument	Original maturity	Market value
31/01/2017	XX1234567	DE	210 (Debt instrument)	2 (Short term)	2200
31/01/2017	XX9876543	DE	317 (Equity instrument)	0 (not applicable)	3000

Mapping metadata using SMCube

SHS

S_121

BSI

FINREP

1100

x10

REF_BSI

S121

REF_FINREP

REF_SHS

AnaCredit

Non-reference
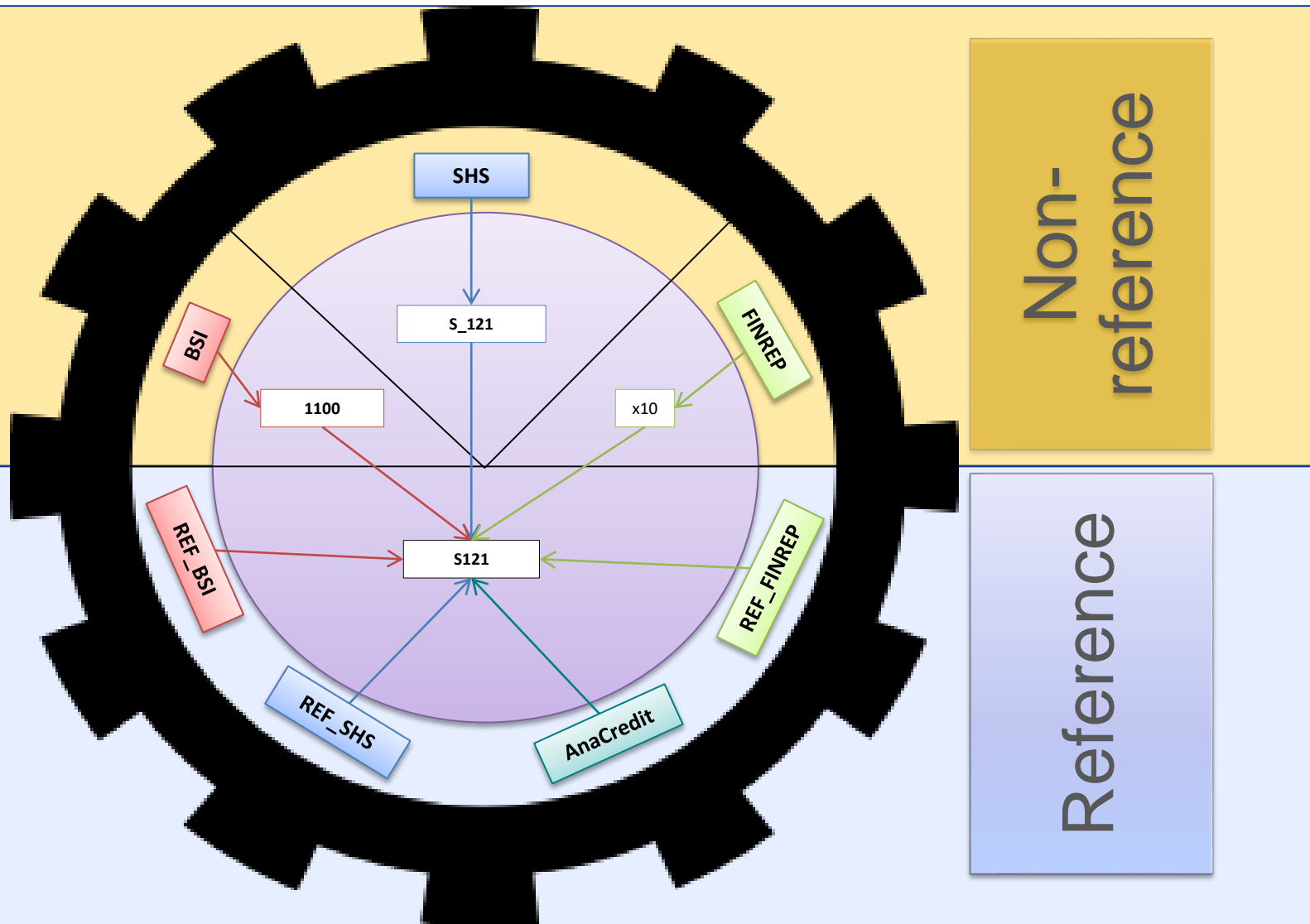
Reference

Example of mapping for data integration BSI/AnaCredit

Mappings in the Single Data Dictionary (SDD)

- Bridging the gap between non-reference and reference codification systems

BSI / MIR non-reference to reference mappings

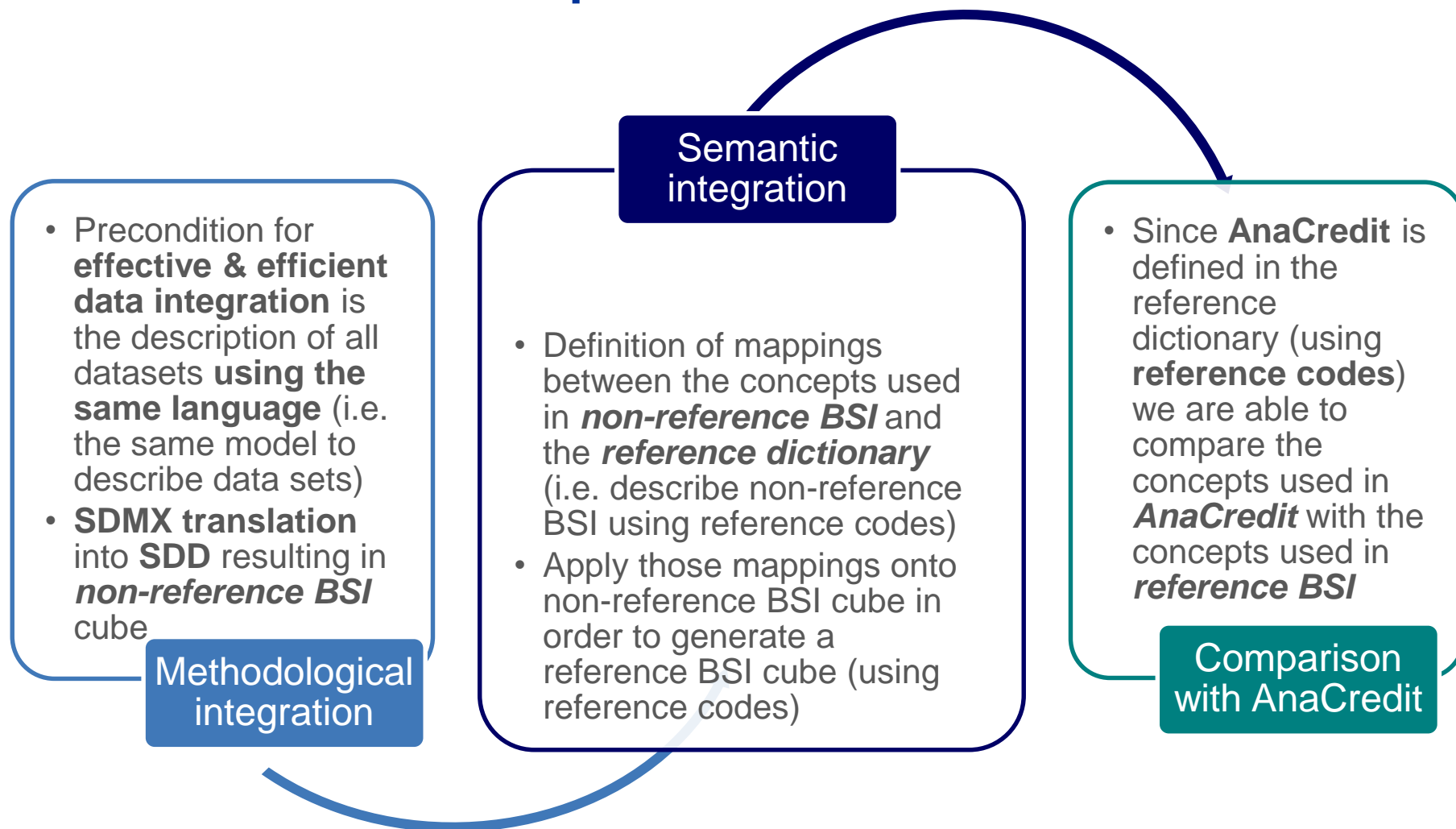- Mapping BSI / MIR concepts into the reference dictionary

Mappings – overview

➢ A **mapping in the SDD** represents the **relationship between concepts** (e.g. variables, members) belonging to **different codification systems / dictionaries** (e.g. BSI, MIR, FinRep…)

➢ Mappings are used in order to **map** (so called) **non-reference concepts into reference concepts**

➢ Mappings are defined in abstract (e.g. variable mappings, member mappings) such that they may be **reused** (e.g. BSI mappings may be reused for MIR)
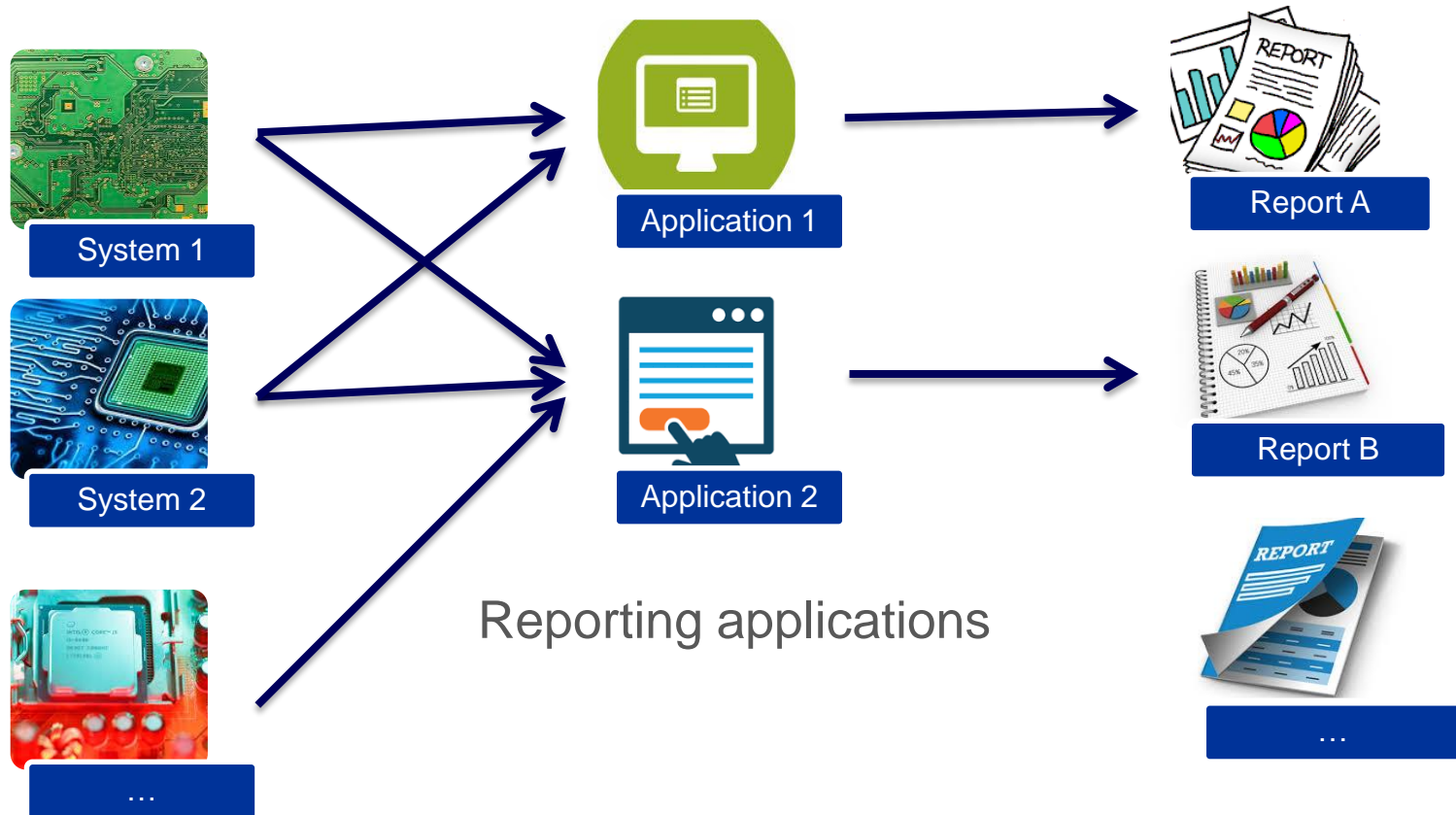
Why?

➢ Concepts that are mapped into the ***reference dictionary*** (i.e. described using reference codes) are ***comparable*** and allow to join data sets using the same concepts (e.g. the same column with the same allowed members) → ***Integration (based on Meta data)***
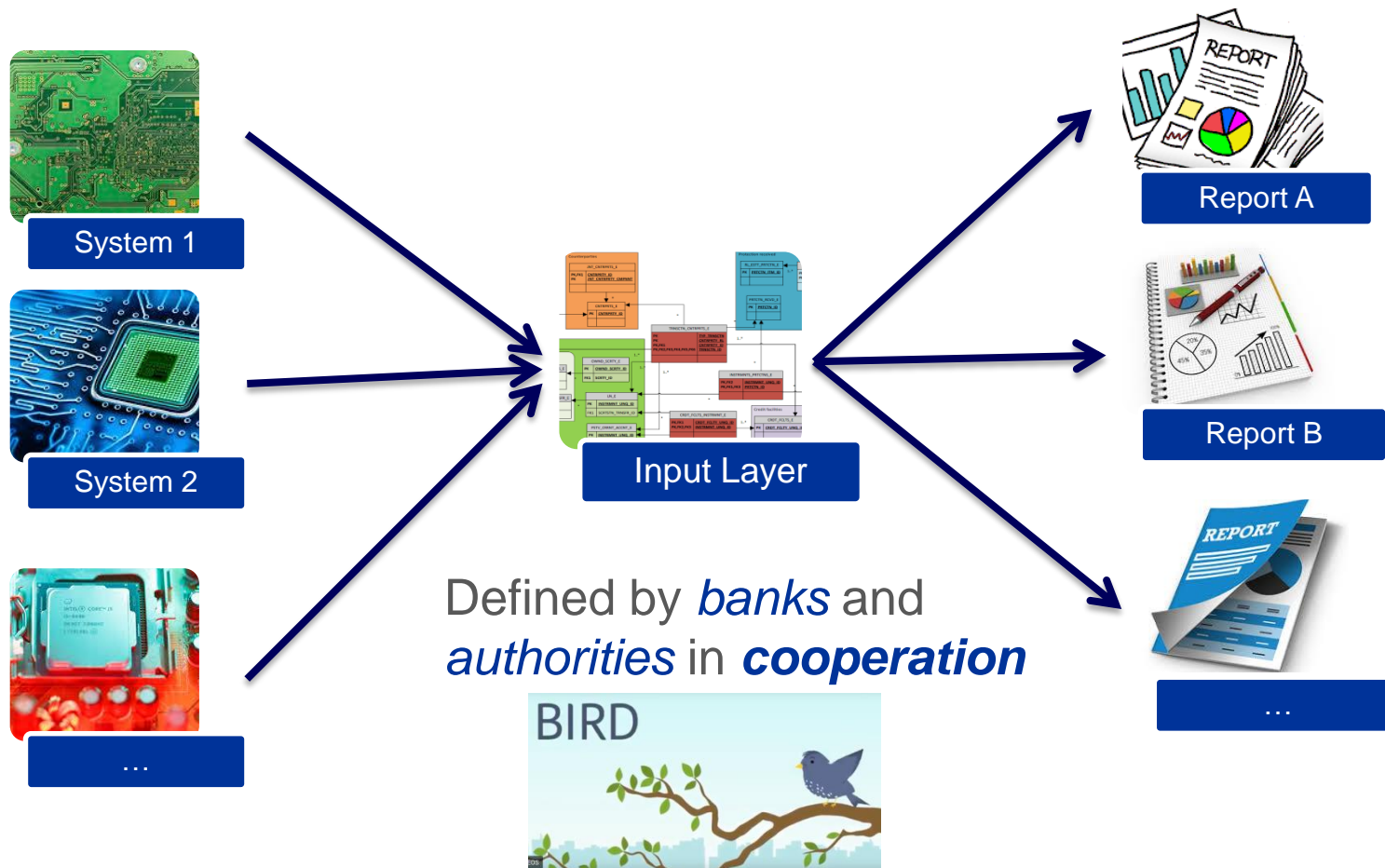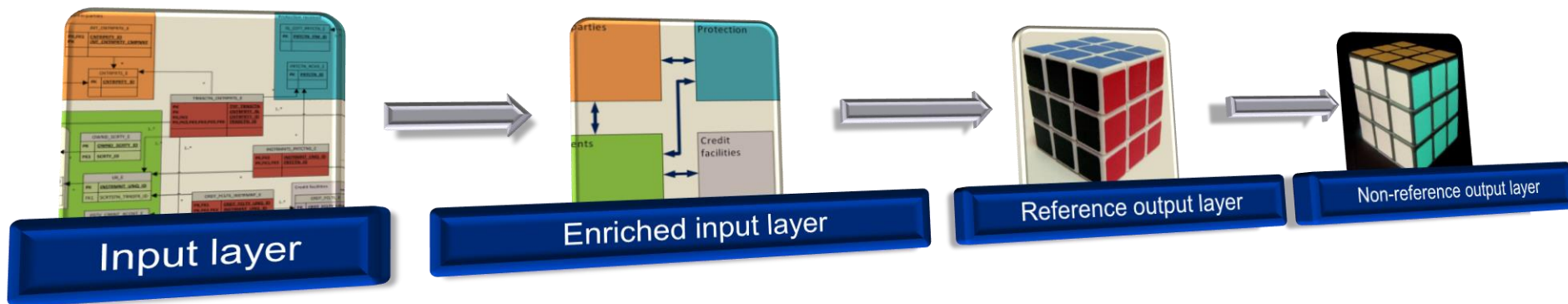
BSI / MIR in the SDD – procedure

Semantic integration

Methodological integration

- Precondition for **effective & efficient data integration** is the description of all datasets **using the same language** (i.e. the same model to describe data sets)
- **SDMX translation** into **SDD** resulting in *non-reference BSI* cube

- Definition of mappings between the concepts used in *non-reference BSI* and the *reference dictionary* (i.e. describe non-reference BSI using reference codes)
- Apply those mappings onto non-reference BSI cube in order to generate a reference BSI cube (using reference codes)

Comparison with AnaCredit

- Since **AnaCredit** is defined in the reference dictionary (using **reference codes**) we are able to compare the concepts used in *AnaCredit* with the concepts used in *reference BSI*

Banks' process to produce reports – the non-integrated

System 1

System 2

…

Application 1

Application 2

Reporting applications

Report A

Report B

…

Banks' process to produce report the integrated approach

System 1

System 2

…

Input Layer

Defined by *banks* and *authorities* in **cooperation**

Report A

Report B

…

Process overview

Input layer → Enriched input layer → Reference output layer → Non-reference output layer

- ➢ Feeding the **Input Layer** from bank's internal IT systems, following the structure of the input cubes defined by BIRD

- ➢ Creation of **Enriched Input Layer** (i.e. an intermediate layer that is used to generate all reporting requirements)

- ➢ Generation of **Reference Output Layer**

- ➢ Generation of **Non-reference Output Layer** (by applying mappings)

The Input Layer
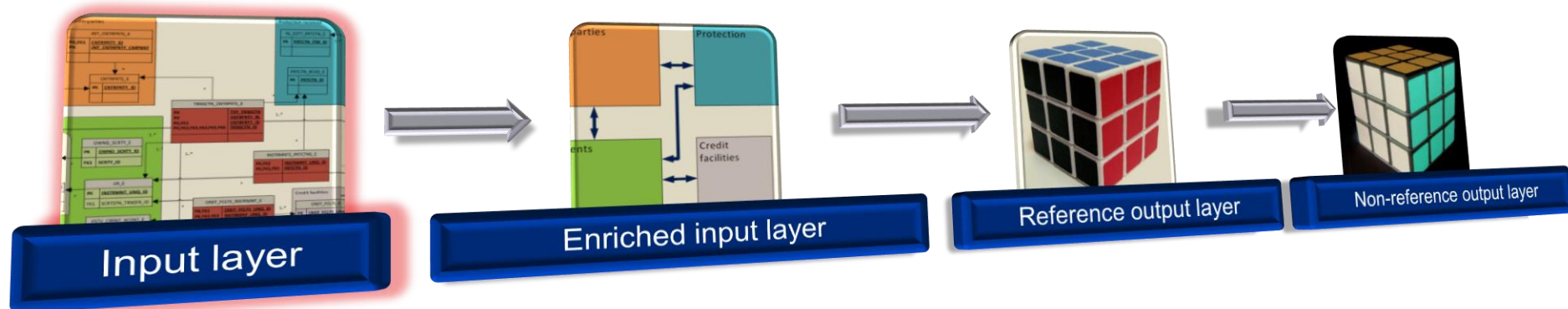
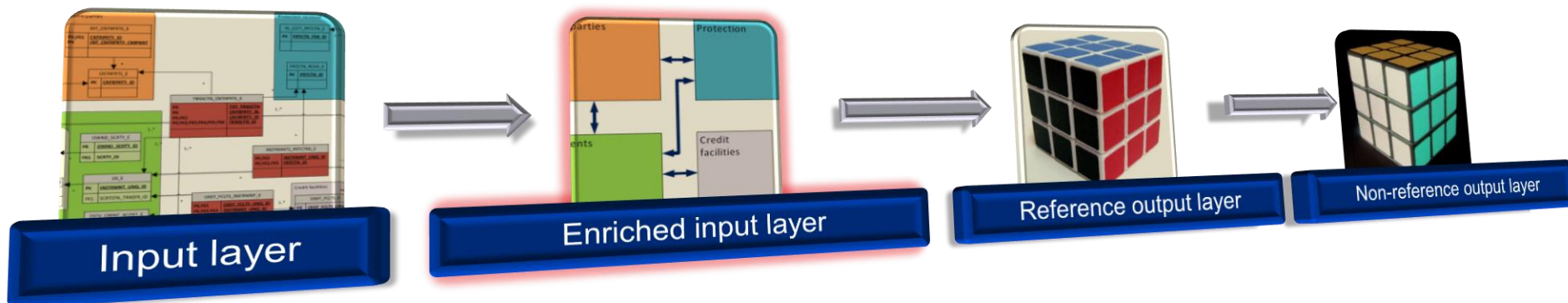What?

➢ **Information** that needs to be collected by **reporting agents** from their internal systems

How? Description?

➢ *Technical guidelines*, providing a **general overview** and **technical instructions for the population** of the Input Layer

➢ *Data dictionary*, **definition of data sets** and their **relationships**

The Enriched Input Layer

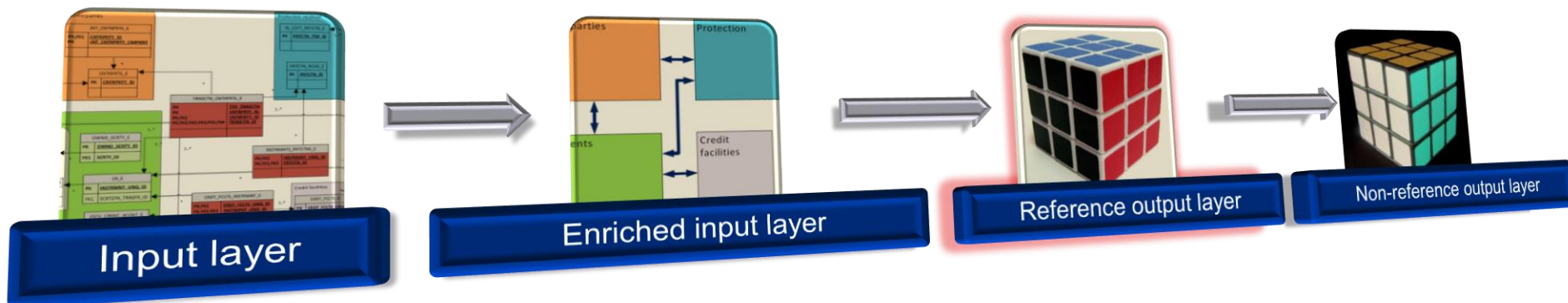
What?

➢ Second layer of data

How?

➢ *Data dictionary*, **definition of data sets** and their **relationships**

➢ *Transformations* based on the **Input Layer**

Why?

➢ **Reuse of concepts** used in **multiple output layers**

The Reference output layer

Input layer → Enriched input layer → Reference output layer → Non-reference output layer

What?

➢ Description of the **Non-reference output layer** (content) using the **reference codification**

How?

➢ *Data dictionary*, **definition of data sets** and **their relationships**

➢ *Transformations* based on **the Enriched input layer**

Why?

➢ **Different codification systems**

The Non-reference output layer
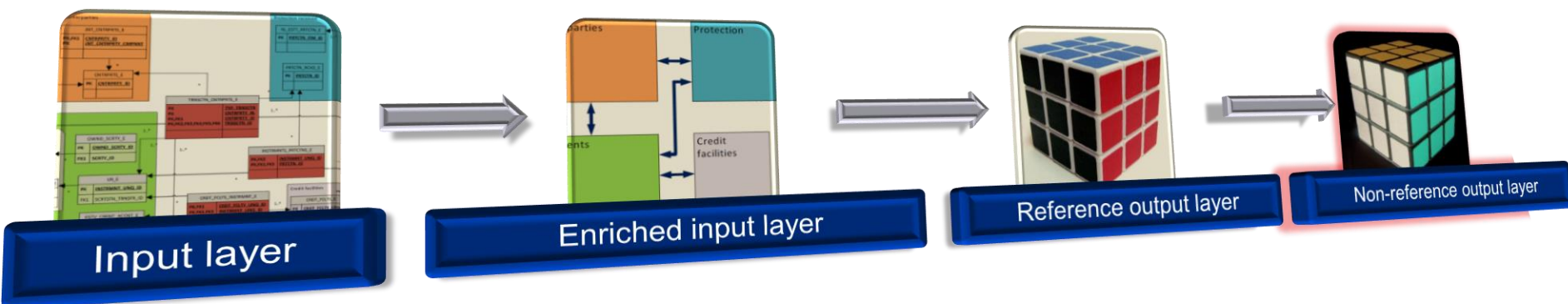
What?

➤ **Secondary reporting requirements** ("data to be reported")

How?

➤ *Data dictionary*, **definition of data sets** and their **relationships**

➤ *Mappings* describing the relationship between the **Reference** and the **Non-reference output layer**   **link to the ERM model**

Bank's Integrated Reporting Dictionary (BIRD)

BIRD

Purpose

- ✓ *Streamline overall process* of reporting from banks to national & European authorities
- ✓ *Improve consistency & quality* of information reported by banks

Benefits

Commercial banks

- ☐ **Data consistency**: Reports produced from a *single input layer* by applying *harmonised transformations*
- ☐ **Lower reporting burden**: *Decrease time & efforts* to analyse & prepare new reports by *exploiting joint knowledge & experience*
- ☐ **Increase efficiency**: Easiness in managing data, satisfying (new) reporting

EUROPEAN CENTRAL BANK
EUROSYSTEM

EBA EUROPEAN BANKING AUTHORITY

Authorities

- ☐ **Formalised, unique, shared** *description* of requirements to banks
- ☐ **Awareness** of reporting agents' *cost drivers*
- ☐ **Fine tune** *the reporting* & better shape/automate data processing

Bank's Integrated Reporting Dictionary (BIRD)

Governance

- BIRD Steering Group (April 2018, December 2018)
- BIRD Expert Group
- BIRD annual workshop with software companies

STC Dialogue with the Banking Industry

Participants

- ✓ **Authorities:**
  NCBs/NCAs, ECB DG/Statistics, Single Resolution Board, EBA

  23%

- ✓ **Banking Industry:**
  European Banking Federation, Commercial Banks

  77%

- Available as a "*public good*" to banks and interested parties: BIRD

BIRD Latest developments

☐ **(Current) *Coverage (BIRD 3.0)***

- Analytical Credit reporting (AnaCredit)
- Securities holdings statistics Group reporting (SHSG)
- Financial Reporting (FinRep 2.7)
- Resolution Planning

☐ ***Work on-going***

- Asset Encumbrance (AE)
- Common Reporting – Credit Risk (CoRep – CR)
- Work stream on data modelling
- Work stream on testing

☐ ***Work plan 2019-2020***

- Securitisation framework DPM 2.9
- FinRep non performing loans DPM 2.9
- Liquidity Coverage Ratio
- Additional Monitoring Metrics for Liquidity

New Database

- **62** input cubes
- **767** transformation schemes