BANCA D'ITALIA
EUROSISTEMA

Integration:
a key-word for success
in central banks' statistics

II FIF Meeting – Lisbon, 4-7 May 2016

Grazia Marchese
Statistical Data Collection and Processing Directorate

BANCA D'ITALIA
EUROSISTEMA

Outline

- Introduction
- Non-integrated versus integrated information systems
- The components of an integrated info sys and their features:
  - Collection
  - Data-warehouse
  - Dissemination
- Governance
- Concluding remarks

BANCA D'ITALIA
EUROSISTEMA

The starting point

"When you cannot measure, ... your knowledge is of a meagre and unsatisfactory kind."

Lord Kelvin

BUT

"When you have wrong measures,… your knowledge is false and misleading."

G. Marchese

ECB President
Mario Draghi

"Policy-making and, indeed, decision-making are only as good as the information on which they are based"

Seventh ECB Statistics Conference, Frankfurt am Main,
15 October 2014

https://www.ecb.europa.eu/press/key/date/2014/html/sp141015.en.html

BANCA D'ITALIA
EUROSISTEMA

Challenges for central banks' statistics

Many functions for a central bank ⇒ many information needs, for:

- Monetary policy
- Macro-prudential policy
- Micro-prudential supervision
- Payment system oversight

A large part of the information needs (and related policies) insist on the same population of reporting agents (banks and other financial intermediaries)

⇓

Risk of overburdening

Central Banks are self-users of their statistical production

⇓

need for accountability

BANCA D'ITALIA
EUROSISTEMA

Challenges for central banks' statistics

New and increasing information needs after the crisis:

- **Need to complete the statistical coverage of** financial sectors, markets and instruments; shadow banking
- **Need to capture heterogeneity and interconnectedness (cross countries, sectors, instruments, markets, intermediaries**
- **More granular data is needed to serve multiple purposes**
- **More flexibility is needed to serve ad-hoc and urgent data requests**

BANCA D'ITALIA
EUROSISTEMA

Challenges:

BEING EFFECTIVE =

Producing "fit for purpose" statistics (comprehensive, time-to-market, high quality…)

AND

BEING EFFICIENT (no entropy, making best use of existing sources, minimizing costs…)

AND

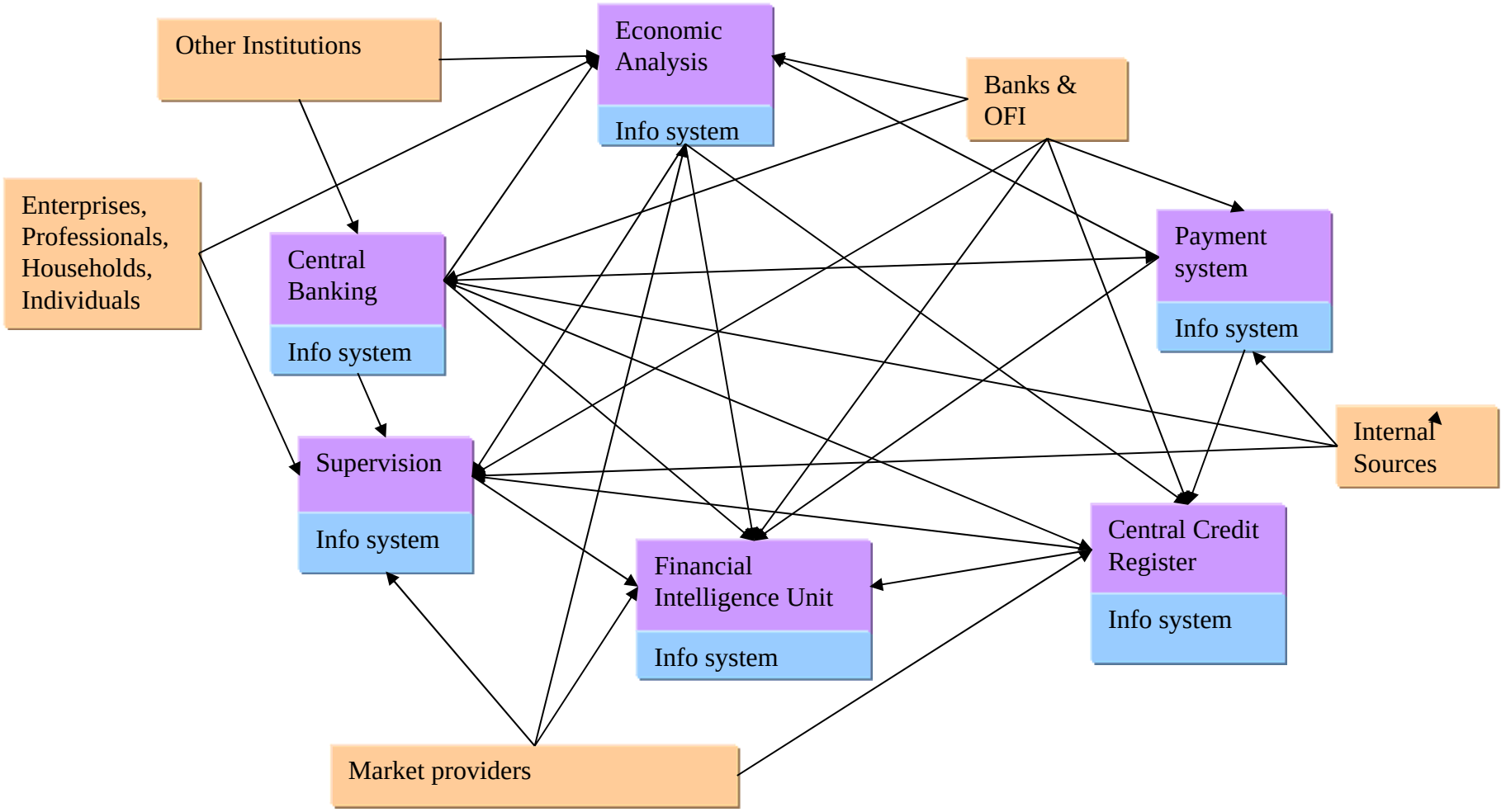
BEING ACCOUNTABLE

Tool:

An integrated model for (each phase of) the statistical production process

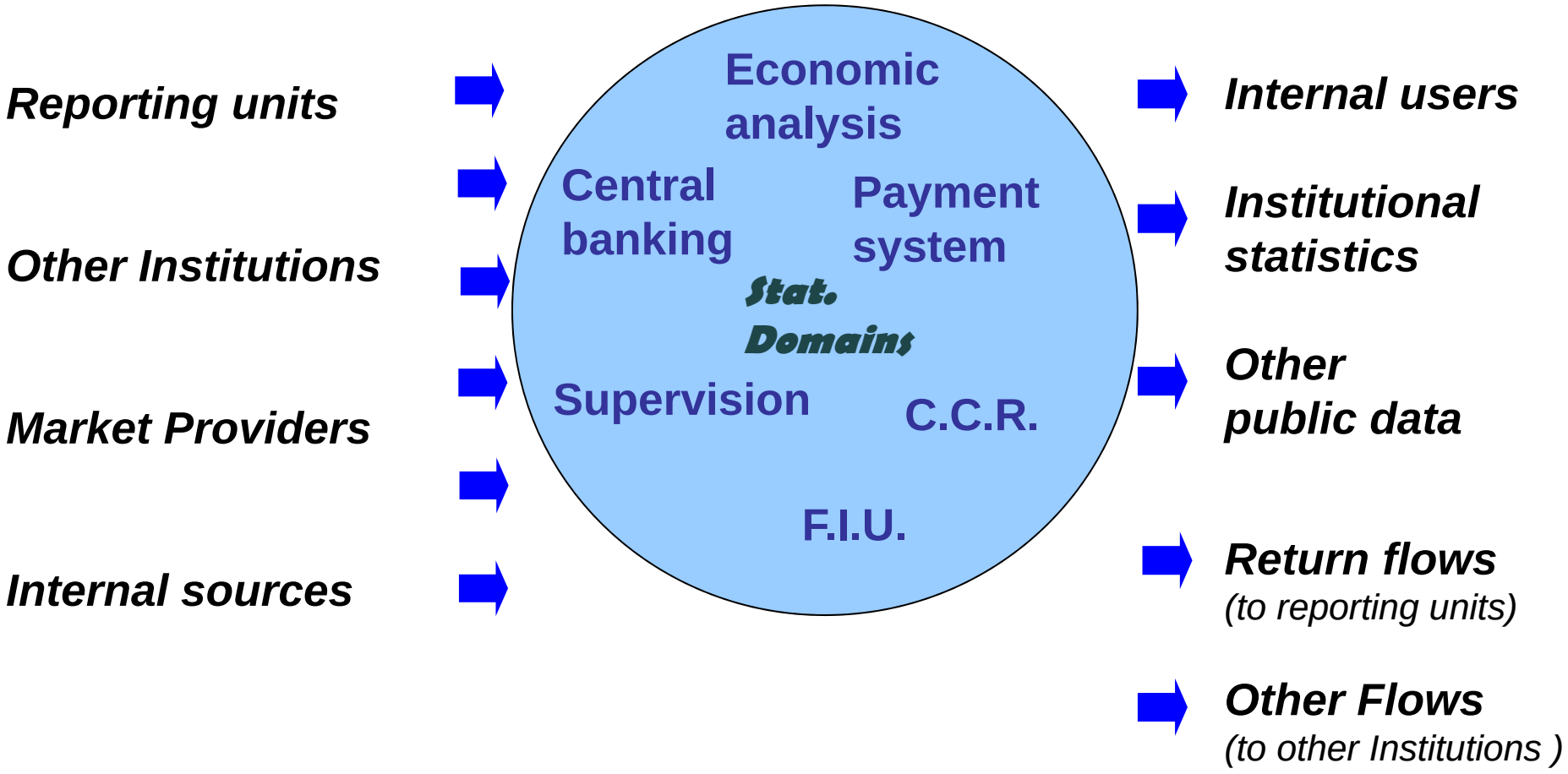
BANCA D'ITALIA
EUROSISTEMA

Non integrated Information System – entropy…..



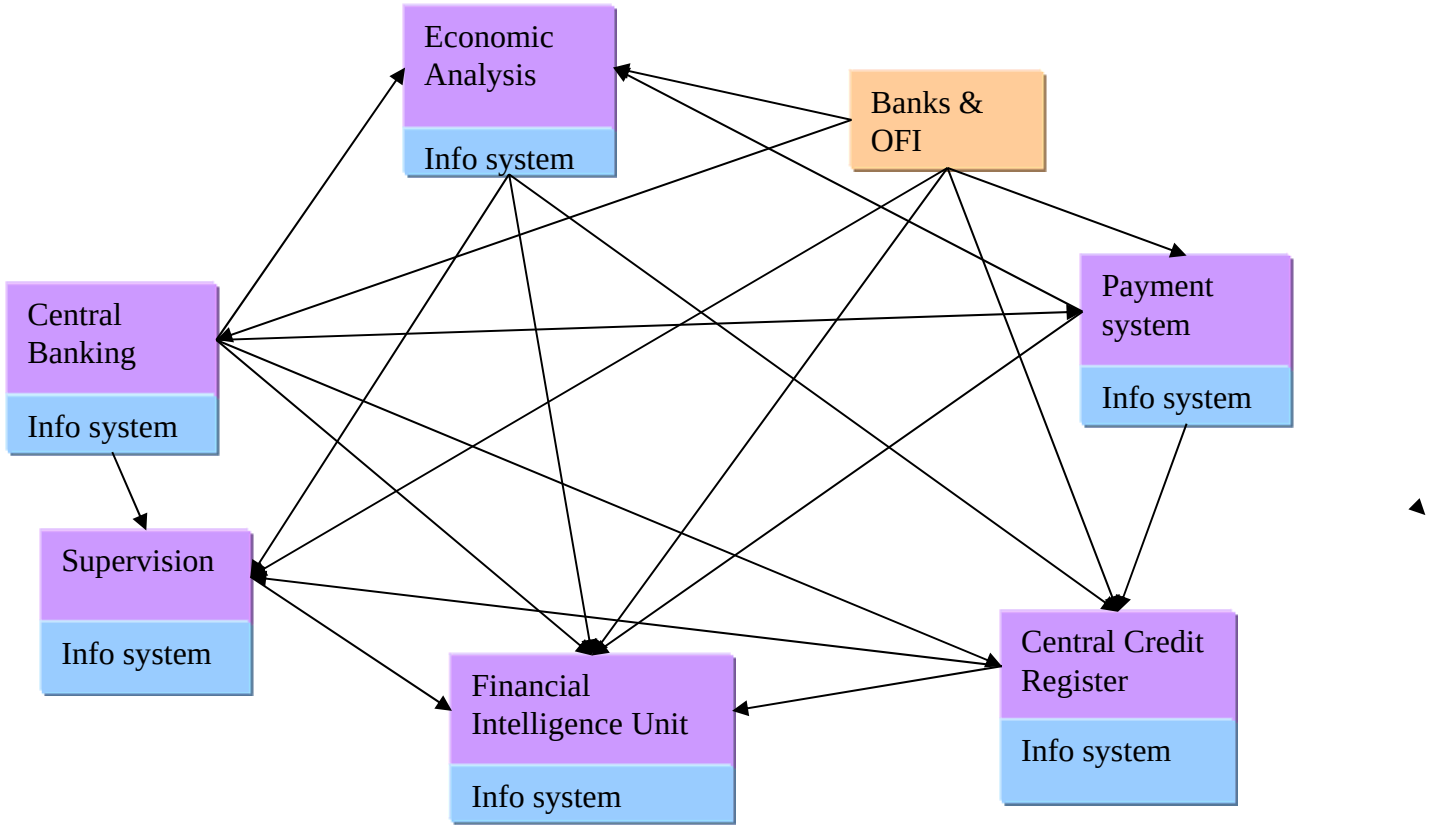
BANCA D'ITALIA
EUROSISTEMA

Integrated Information System – streamlined!



BANCA D'ITALIA
EUROSISTEMA

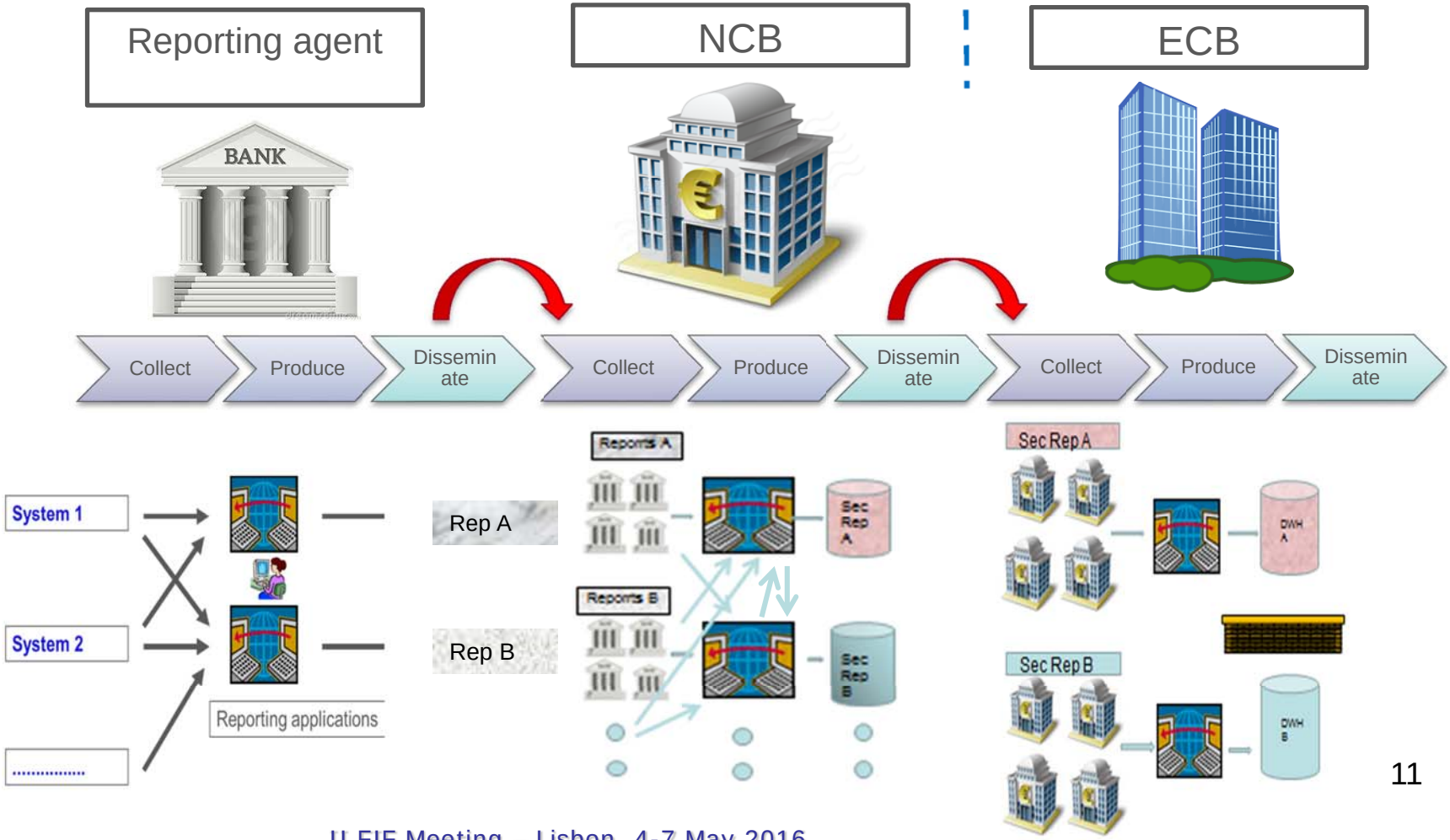
Non integrated Information System (subset)



BANCA D'ITALIA
EUROSISTEMA

Non-integrated Information System (subset)

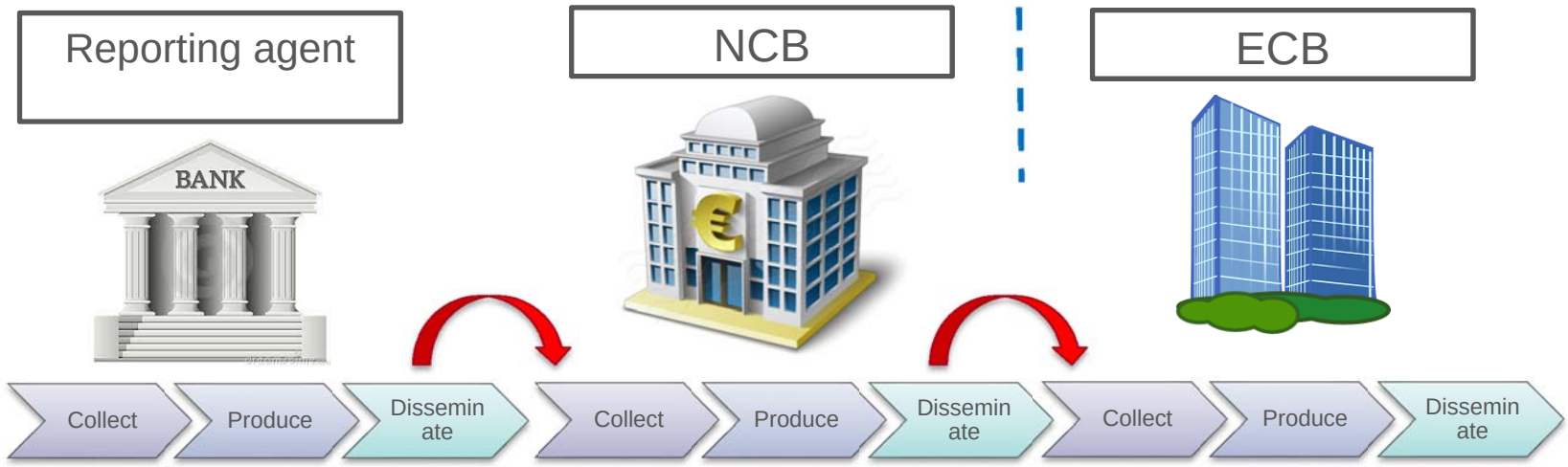
for whole process : entropy + silos

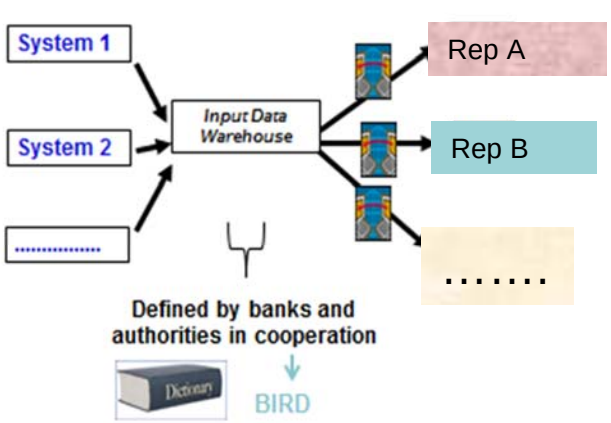


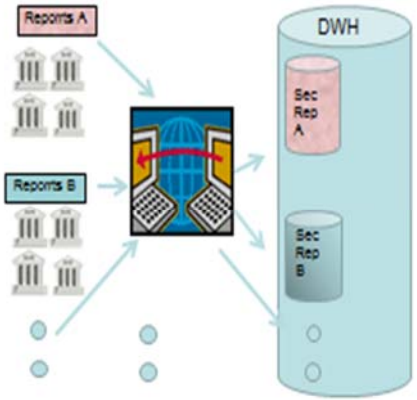
BANCA D'ITALIA
EUROSISTEMA

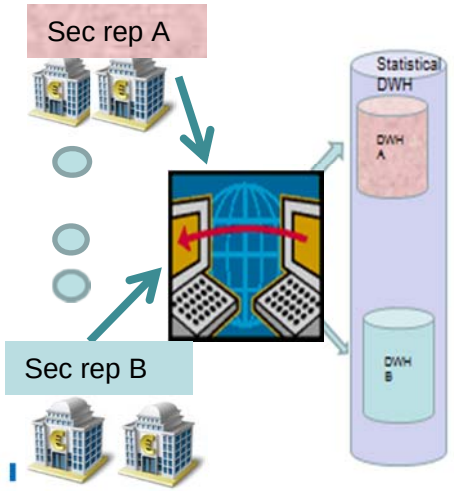
Integrated Information System (subset)

for whole process : no duplication, no redundancy, no silos





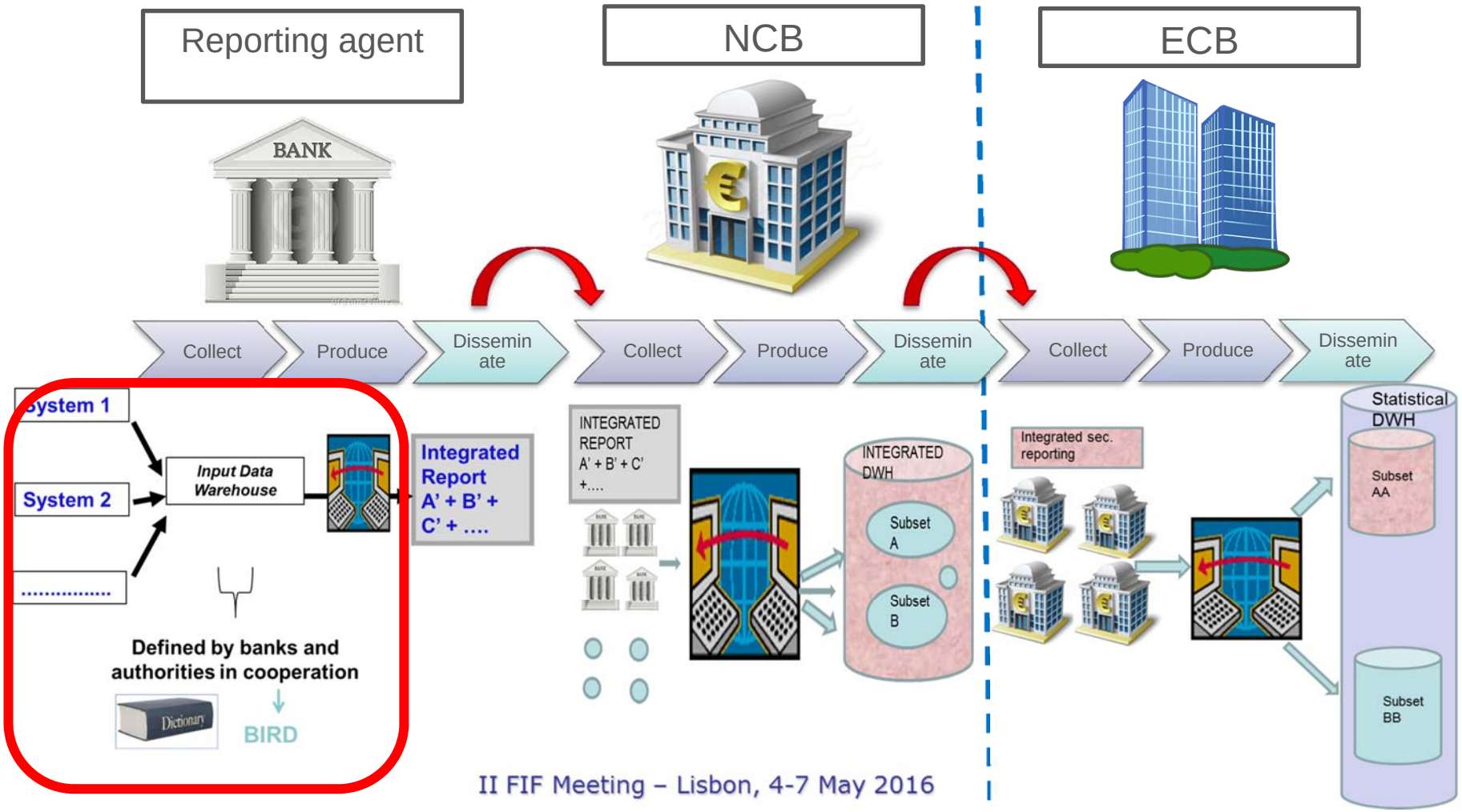




BANCA D'ITALIA
EUROSISTEMA

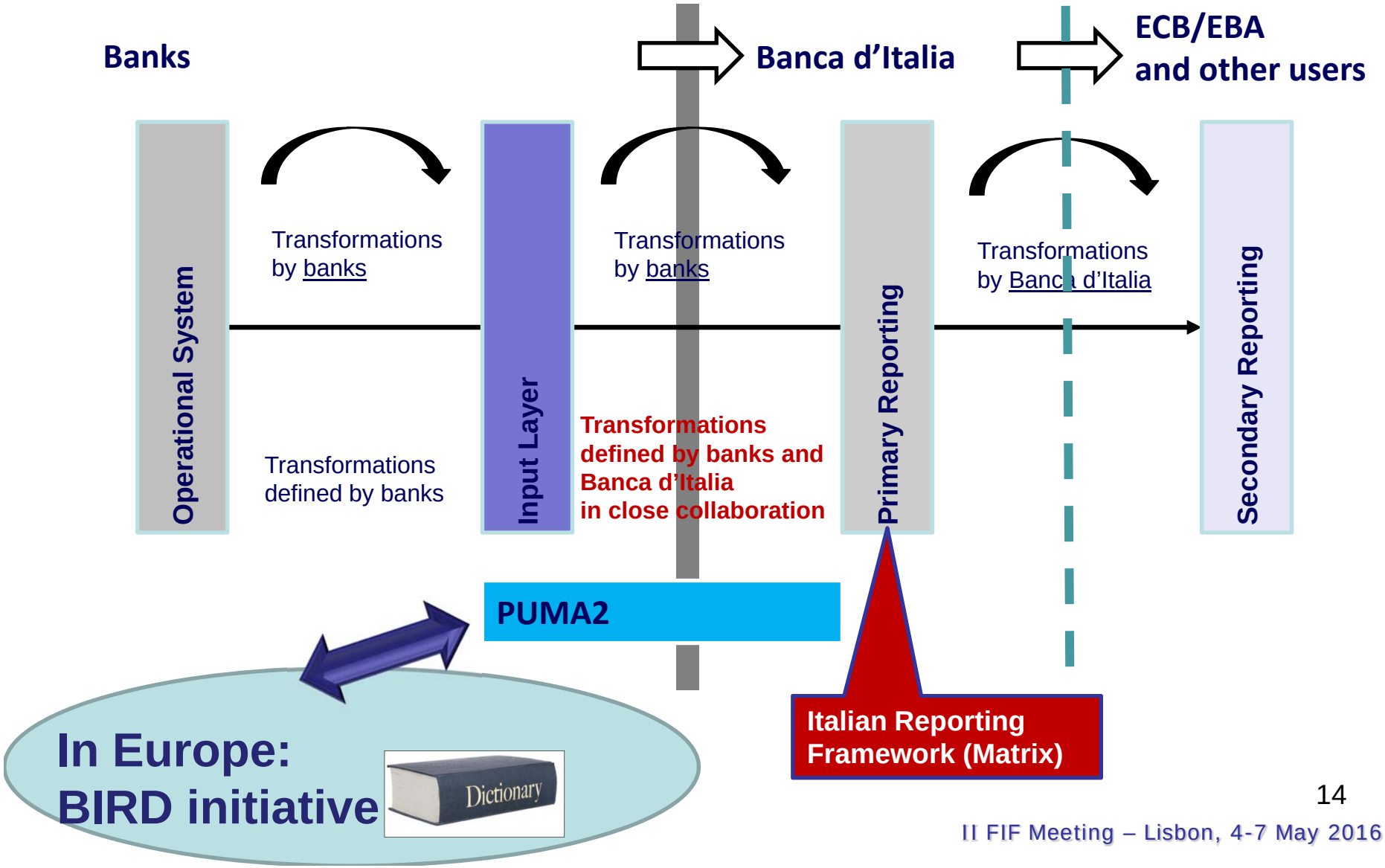
Integrated Information System (subset)

for whole process : no duplication, no redundancy, no silos



BANCA D'ITALIA
EUROSISTEMA

Integration – for reporting agents (banks)



BANCA D'ITALIA
EUROSISTEMA

PUMA BIRD: features

What:

- Precise definition of input data to be extracted from banks' operational databases: the «**input layer**»
- Data **transformation** to be applied to data in the input layer to derive reports requested by authorities

How:

- Deliverables are produced and maintained by reporting agents and authorities in close **cooperation** and on a **voluntary basis**
- Deliverables are a «**public good**», to reporting agents and all interested parties
- **Adoption by reporting agents is voluntary**

BANCA D'ITALIA
EUROSISTEMA

PUMA ≈ BIRD: advantages

For the banks:

- Data consistency: integrated production of different reports
- Genuine and univocal interpretation of requirements
- Contained reporting burden
- Facilitates IT solution for reporting

For the authorities:

- Higher data quality
- Traceability of the process
- Flexibility: quickness to satisfy new requirements

For both:

- Cost/benefit ratio is minimised
- Better information → better policy decisions → higher returns for the society as a whole → higher credibility of CB → higher effectiveness of policies

BANCA D'ITALIA
EUROSISTEMA

Integration in primary reporting

Integrated Information System (subset) for whole process...

- Advantages are maximized and costs minimized when reports for different policy needs are integrated into a single framework that is complete, internally consistent and non-redundant.
- Completeness, internal consistency, no redundancy are conceptual requirements for the overall framework; splitting the framework into subsets to be delivered at different deadlines and with different frequencies is always possible.
- An integrated reporting framework requires a certain level of granularity.
- A window of opportunity is given by the new awareness by the authorities that more granular data are needed to better preserve financial stability.

BANCA D'ITALIA
EUROSISTEMA

Integration in primary reporting

In Italy:

the «Matrix of bank accounts»: long lasting approach, dating back to mid '70s

In Europe:

The European Reporting Framework (ERF); the project is at an early stage and currently under conceptual development.

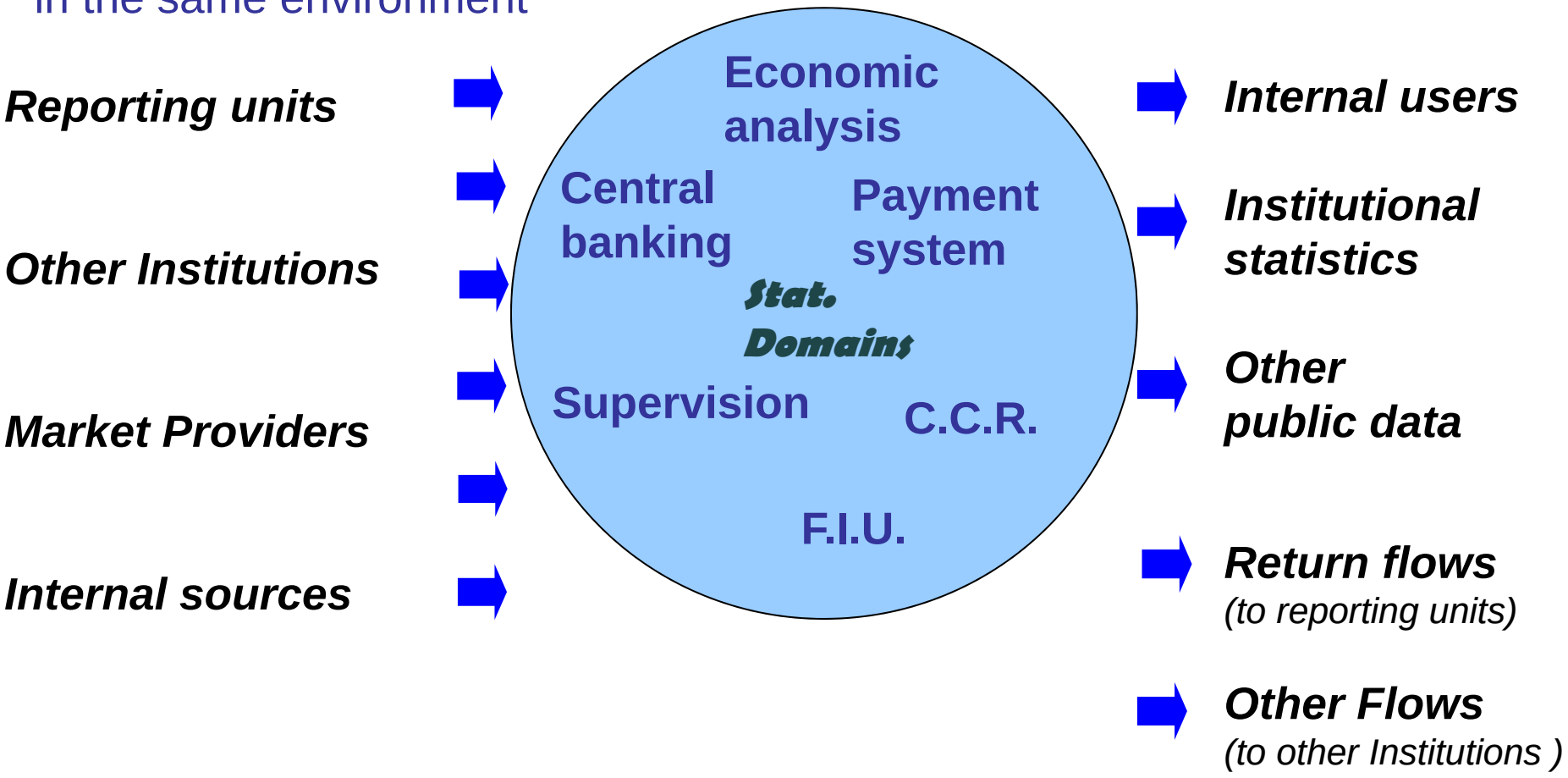
In the EU context, the eventual implementation of the ERF would add a further dimension to the benefits from an integrated primary reporting: cross-country standardization in banks' reports would better ensure level playing field and would facilitate financial integration.

Integrated Information System (subset) for whole process...

BANCA D'ITALIA
EUROSISTEMA

Integrated Data-warehouse

I-DWH is more than simply a collection of datasets that are stored together in the same environment



BANCA D'ITALIA
EUROSISTEMA

Integrated data-warehouse

For an I-DWH a comprehensive conceptual framework is needed.

Ingredients:

- A reference information model (how information is structured)
- A data dictionary (internally coherent and non-redundant set of attributes describing the DB content ⇨ domains, dimensions, code-lists, measures, attributes ⇨ metadata)
  - ★ the data dictionary includes DERIVED VARIABLES with their metadata, among which:
    - ✦ hierarchical and algebraic links (to derive them from elementary or other derived variables

- Cross-time correlation (changes of definitions over time)
- Versioning (variations, updates)

BANCA D'ITALIA
EUROSISTEMA

Integrated Data-warehouse

A **METADATA-DRIVEN DWH**:

Expression language (EXL) is used to obtain derived variables ⟹ definitions can be made automatically

★ Very powerful command over DWH content:

– Live definitions (integrity) ⟹ easy life for data administrators and users

– Quality checks defined with the same language used for derived variables; results can be stored into the DWH for further processing and inquiring

BANCA D'ITALIA
EUROSISTEMA

Integrated Data-warehouse

Main components

- Primary data sources
  - Proprietary
  - External

  } → statistical data, administrative data
- Derived statistics

Data kinds

- Multidimensional
- Time series
- Cross section

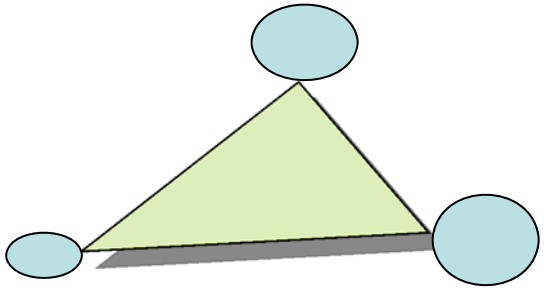
BANCA D'ITALIA
EUROSISTEMA

Integrated Data-warehouse

Main components

- Macro data
- Micro data
  - CCR
  - Enterprise bal. sheets
  - Security holdings
  - BOP direct reporting
  - Households survey
  - ….
- Registers
  - Entity register
  - Security register

PIVOTAL ROLE



BANCA D'ITALIA
EUROSISTEMA

The role of registers

Backbone for the statistical production process:

- increase quality and coherence of statistics
- empower granular data collection
- minimise burden on reporting agents
- Key for linking granular data from different domains

Harmonization: **unique BdI** identification code + multiple internal and external identifiers

Unique **international** identifier: ISIN code

«Certified» reference data on entities and securities used in statistical collection, production and dissemination *(e.g.: Securities Issues, Securities Holdings, IF, FVC and MFI Balance sheets, Central Credit Register)*

BANCA D'ITALIA
EUROSISTEMA

Integrated Data-warehouse: recap

- Data stored and processed **avoiding redundancies**
- **Integrated use** of data from different sources and **reuse** for many purposes
- The evolution of statistics is governed in a **harmonized** and **consistent** way across the wide range of internal uses
- Users can exploit the information collected and elaborated with a clear **understanding** about data contents
- Information is **shared** among many functions
- Data are **accessible to all users** who have rights: **confidentiality** constraints and information circularity in general **can be managed** in a orderly way
- Integration **is not integralism**: a few «statistical communities» exist, to allow for autonomy and flexibility when they are an absolute must.

BANCA D'ITALIA
EUROSISTEMA

Integration (= harmonisation) in data dissemination

- Especially among producers of official statistics, their DWH is an open system: output for dissemination may be input for others, and vice-versa

- A common "language" is needed to describe and exchange statistical output

- standardized classifications and code-lists ⟷ sdmx Statistical Data and Metadata eXchange

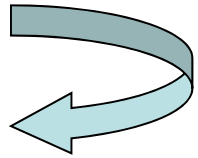
- standardized taxonomies ⟷ XBRL eXtensible Business Reporting Language

BANCA D'ITALIA
EUROSISTEMA

Integration (= harmonisation) in data dissemination

Challenges:

- Ideally only globally harmonised and shared classifications and code-lists should be used in the statistical data dictionary; in practice existing standardized classif./codelist do not yet cover the whole set of phenomena of interest for each agency
- Harmonised and proprietary classif./codes have to coexist and be compatible with each other ⇒ burden for data admin.



Mitigation measures:

- Require standard codes in the reporting schemes, so as to replace proprietary ones in all phases of data cycle: collection, compilation and dissemination

Benefits:

- Using harmonized classif./code-lists in disseminating statistics greatly enhances usability of data and better qualifies them as "public good":
  a "plus" for ACCOUNTABILITY
- Easier data sharing and reuse among producers: COST REDUCTION
- Both for internal and external users: immediate comparability and joint use of different sources ⇒ MORE EFFICIENCY/EFFECTIVENESS

BANCA D'ITALIA
EUROSISTEMA

An integrated Information System needs adequate **GOVERNANCE**:

Integrated management of business information along its entire life cycle (definition of contents, collection, verification, production, analysis, dissemination) has to be ensured

An Internal Committee acts as:

- **Coordinating body for users' requirements**
- **DWH owner**
- **Responsible for change management**

- **Bol's Statistics Committee, chaired by a member of the Board, composed by the Heads of statistics producers and users Directorates (statistics, supervision, monetary operations, payment system, research, etc.)**

BANCA D'ITALIA
EUROSISTEMA

Concluding remarks

Statistics are a public good: it is a duty for us to build this very valuable asset in such a way to ensure that

COST/BENEFIT ratio is minimized

To reduce costs: **integrated/multipurpose data collection**. For banks and OFIs an Integrated reporting Framework (ERF ≈ Italian "MATRIX"), backed by co-operative arrangement with reporting agents to build the (logical - conceptual) bridge between their internal data repositories and primary reporting requirements (BIRD ≈ Italian PUMA). The same ingredients also increase quality of collected statistics.

To maximize benefits: **integrated DWH**, assisted by the pivotal role of registers, to exploit all the synergies from combined uses of the various sources and to better serve information needs for policy decision.

Smart dissemination practices, using harmonized concepts and international standards, to contribute to **accountability** and to progress in research

Cross-country harmonization to amplify the advantages of the "integrated approach"

BANCA D'ITALIA
EUROSISTEMA

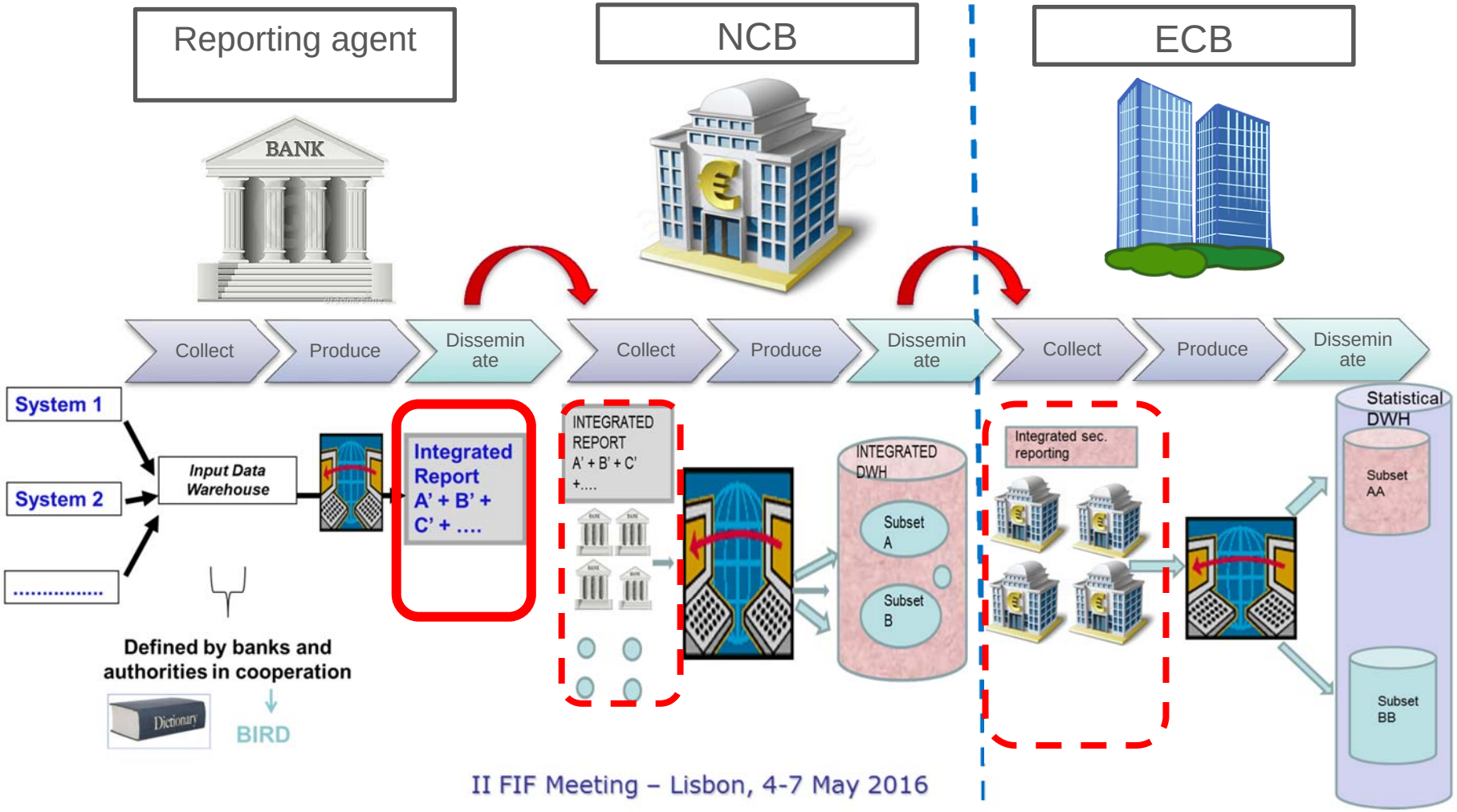
Thank you
for your attention!

Any question?

BANCA D'ITALIA
EUROSISTEMA

Integrated Information System (subset)

for whole process : no duplication, no redundancy, no silos



BANCA D'ITALIA
EUROSISTEMA

Integrated Data-warehouse I-DWH is more than simply a col...

Integrated Information System (subset)

for whole process : no duplication, no redundancy, no silos