NEW STRATEGY OF DATA SHARING AND DATA ACCESS IN STATISTICS

THE VIEW FROM BANCO DE PORTUGAL

OCTOBER 2023

BANCO DE PORTUGAL
EUROSISTEMA

PRESENTATION OVERVIEW

01 | BANCO DE PORTUGAL'S STATISTICS DEPARTMENT

02 | WHERE IT ALL STARTED

03 | FUTURE VISION OF OUR STATISTICAL INFORMATION SYSTEMS

04 | DEFINING A COMMON SET OF CONCEPTS

05 | CLOSING REMARKS

BANCO DE PORTUGAL'S STATISTICS DEPARTMENT

01

BANCO DE PORTUGAL'S STATISTICS DEPARTMENT
ORGANIZATIONAL STRUCTURE

MONETARY AND FINANCIAL STATISTICS	EXTERNAL STATISTICS	CENTRAL BALANCE SHEET	FINANCIAL ACCOUNTS	DATA MANAGEMENT	FINANCIAL SYSTEM MICRODATA	COMMUNICATION AND PLANNING	STATISTICS QUALITY

BANCO DE PORTUGAL'S STATISTICS DEPARTMENT

OUR STATISTICS

CENTRAL BALANCE SHEET STUDIES

INDEBTEDNESS OF THE NON-FINANCIAL SECTOR

EXCHANGE RATES

INCOME AND SAVING

BUDGET OUTTURN

INTEREST RATES

GENERAL GOVERNMENT FINANCIAL ACCOUNTS

SECURITIES

ECONOMIC AND FINANCIAL INDICATORS

INTERNATIONAL RESERVES

PENSION FUNDS

MFI BALANCE SHEET

GENERAL GOVERNMENT FINANCING

BALANCE OF PAYMENTS

EXCHANGE RATES INDICES

INTERNATIONAL INVESTIMENT POSITION

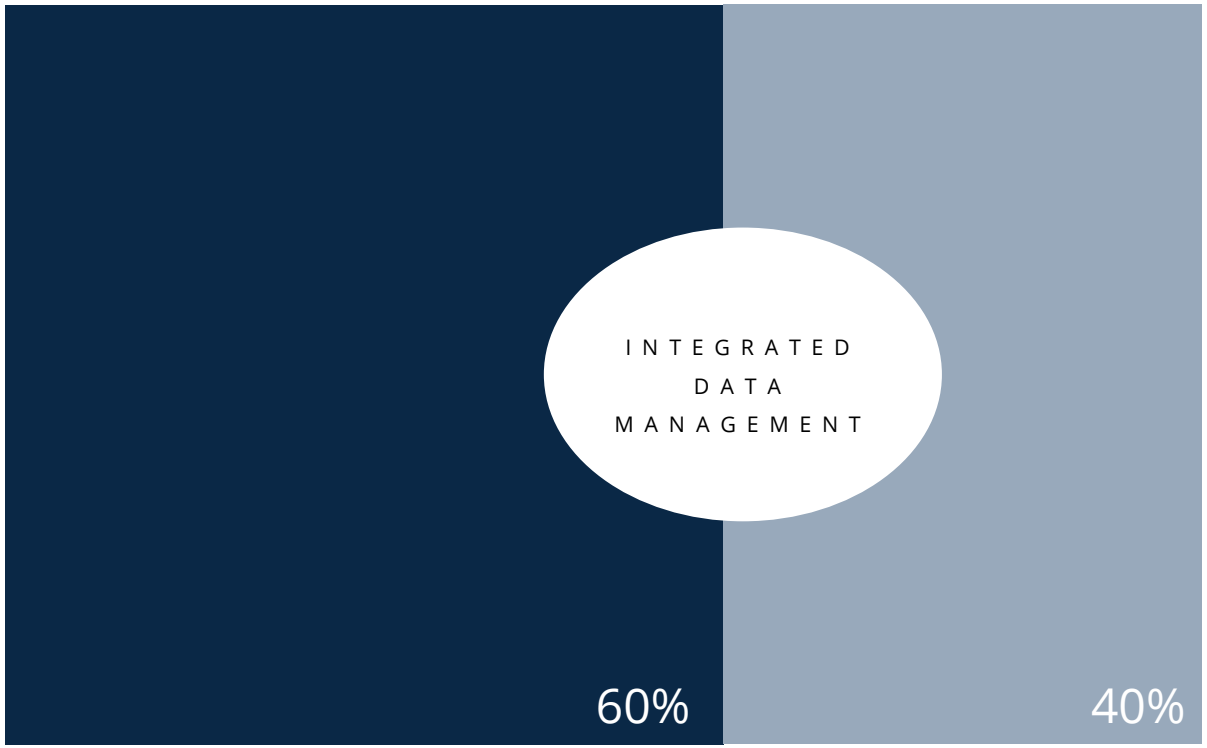
NATIONAL FINANCIAL ACCOUNTS

GENERAL GOVERNMENT DEBT

BANCO DE PORTUGAL'S STATISTICS DEPARTMENT

WHAT DO WE DO?

PRODUCE AND DISSEMINATE STATISTICS



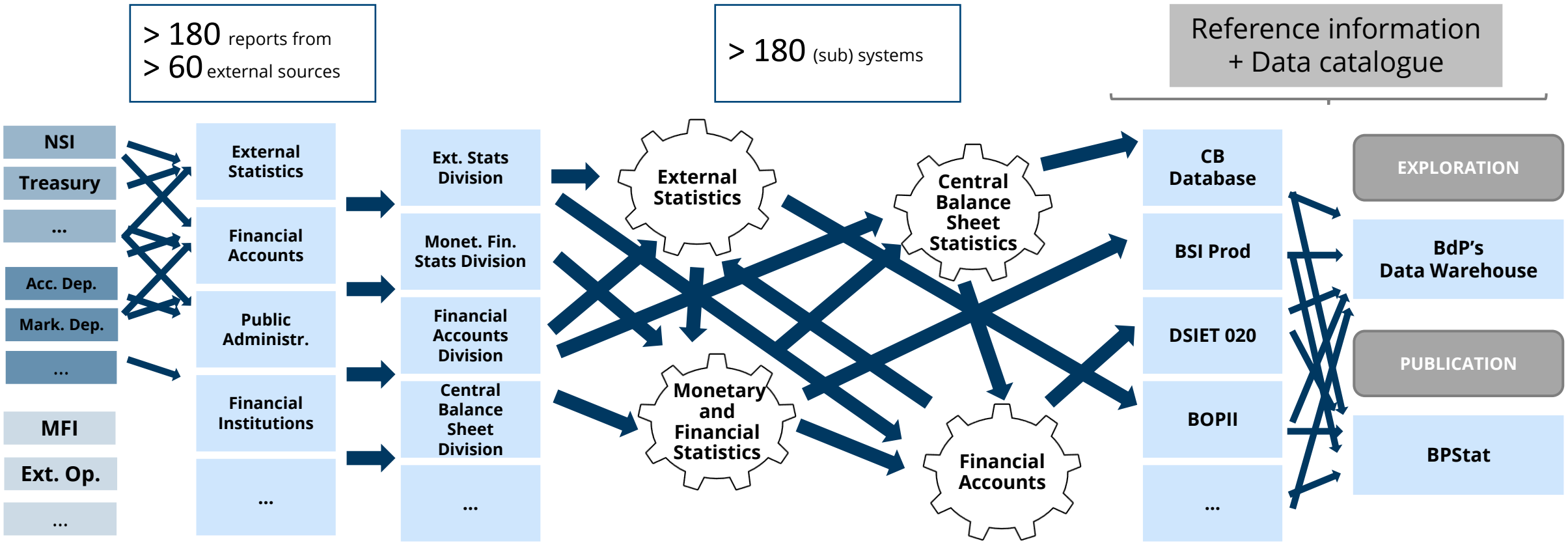
MANAGE OTHER DATABASES

[CENTRAL CREDIT REGISTER, BANKING ACCOUNTS DATABASE, IN-HOUSE CREDIT ASSESSMENT SYSTEM]

WHERE IT ALL STARTED

02

WHERE IT ALL STARTED



COLLECTION

PRODUCTION

Each division/unit is reponsible for the treament of its relevant source

Each domain is responsible for its own quality control procedures

Each system has its own information repository

WHERE IT ALL STARTED

Main thoughts

Each domain was responsible for the treatment of its own relevant set of data.

Adapted to the purpose of each division.

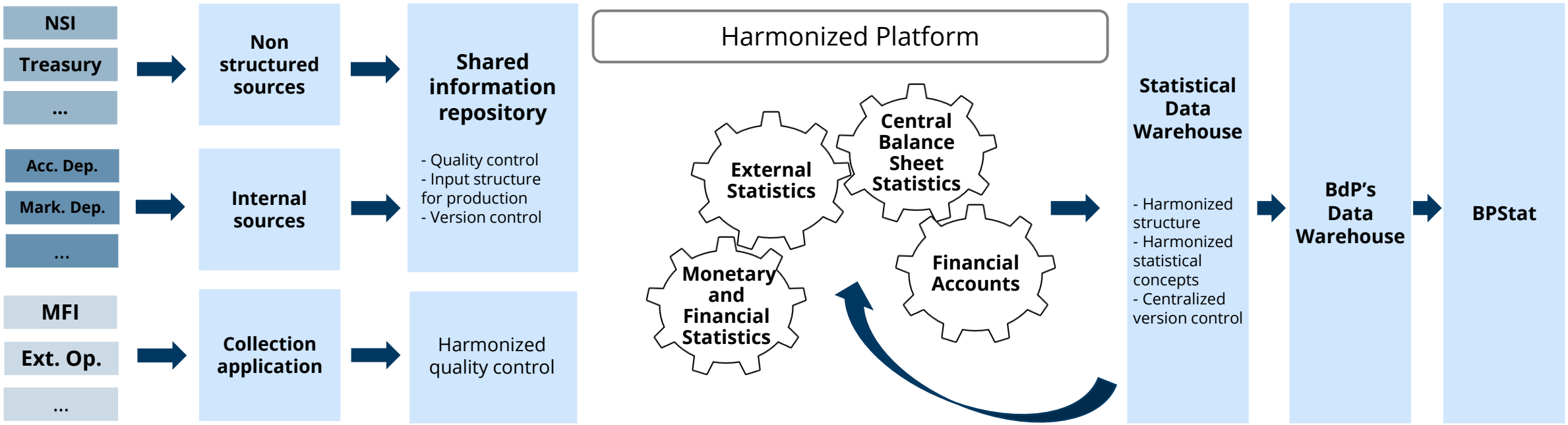
No harmonized concepts, processes or systems.

FUTURE VISION OF OUR STATISTICAL INFORMATION SYSTEMS

03

FUTURE VISION OF OUR STATISTICAL INFORMATION SYSTEMS
UNTANGLING THE WEB

Reference information + Data catalogue

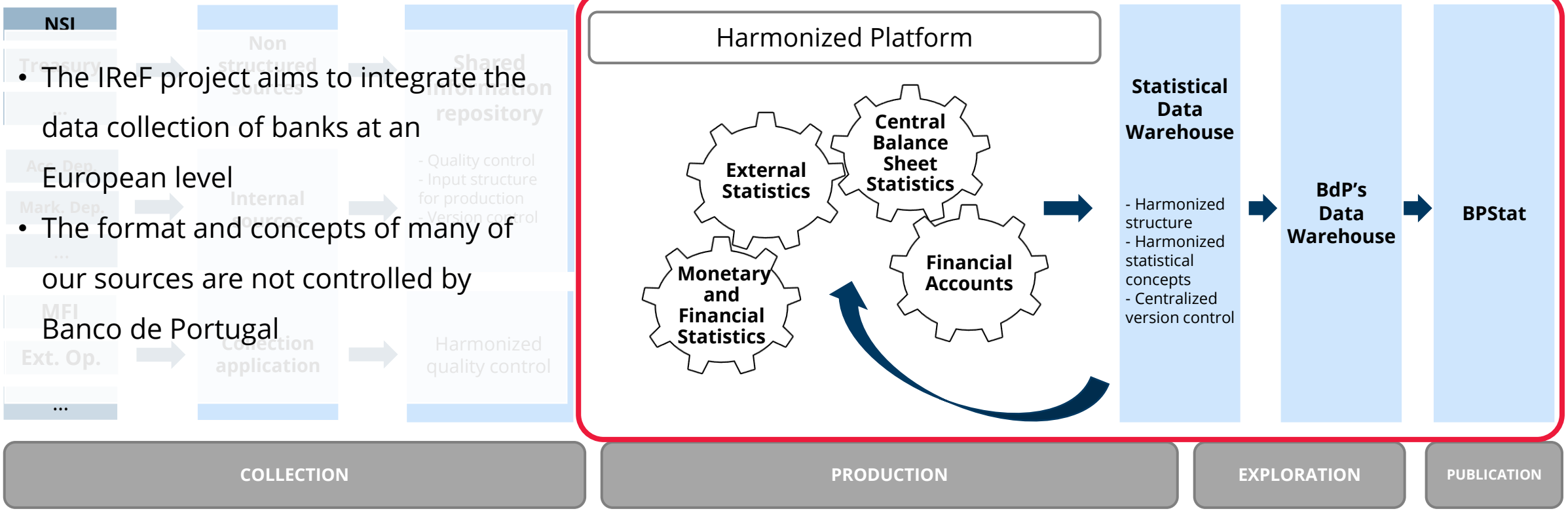


Simplify, Harmonize and Enable.

FUTURE VISION OF OUR STATISTICAL INFORMATION SYSTEMS
UNTANGLING THE WEB

Reference information + Data catalogue

- The IReF project aims to integrate the data collection of banks at an European level
- The format and concepts of many of our sources are not controlled by Banco de Portugal



FUTURE VISION OF OUR STATISTICAL INFORMATION SYSTEMS
UNTANGLING THE WEB

Main thoughts

Simplify and harmonize our domains.

Enhance our teams' knowledge on where and how to intervene.

Homogeneous and consolidated domains, easier to maintain and evolve.

Promote transparency, auditability and efficiency in our processes.

Information sharing, with less adaption costs.

Semantic integration as a cornerstone of our strategy.

DEFINING A COMMON SET OF CONCEPTS

04

DEFINING A COMMON SET OF CONCEPTS

WHY DOES IT MATTER?

ALL DOMAINS “TALK” THE SAME LANGUAGE.

HARMONIZED CONCEPTS FACILITATE COMPARISON.

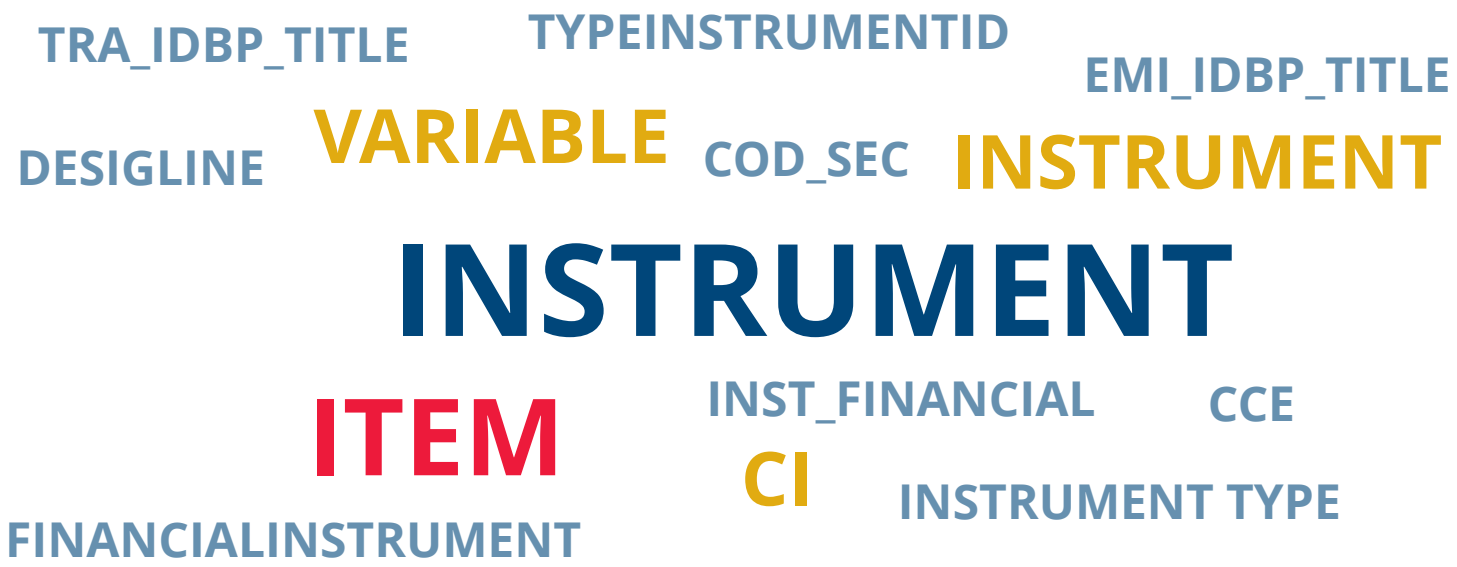
PRODUCTION “COSTS” ARE REDUCED.

DIFFERENT INTERPRETATIONS ARE AVOIDED.

FULL KNOWLEDGE OF ALL THE DATA AVAILABLE IN-HOUSE.

DEFINING A COMMON SET OF CONCEPTS
UNHARMONIZED DATA – THE ISSUE

WHEN WE SEARCH FOR THE DIMENSION **INSTRUMENT**, WHAT DO WE FIND IN THE DATABASES?



DEFINING A COMMON SET OF CONCEPTS
UNHARMONIZED DATA – THE ISSUE

MULTIPLE AND DIFFERENT SOURCES, REGULATIONS AND REPORTS.

CONCEPTS WITH HETEROGENEOUS DESCRIPTIONS.

DIMENSIONS (VARIABLES) WITH MULTIPLE CONCEPTS.

DIFFICULTIES WHEN CROSSING DATA BETWEEN DOMAINS.

DEFINING A COMMON SET OF CONCEPTS
UNHARMONIZED DATA – HOW CAN WE SOLVE IT?

A TASK FORCE TO HARMONIZE CONCEPTS.

DEFINING A COMMON SET OF CONCEPTS
HARMONIZATION STRATEGY

A | DEFINE A GLOSSARY

B | COLLECT STATISTICAL OUTPUTS

C | HARMONIZATION CYCLE

D | RESULTS

DEFINING A COMMON SET OF CONCEPTS

A) DEFINE A GLOSSARY

Investment funds (example)

Original maturity (example)

Over 1 and up to 2 years (example)

PRODUCTION OUTPUT

Production output is a table/database that contains the result of statistical compilation at the most granular level.

DIMENSION

The dimension provides an overview of the statistics produced.

It's composed by:

- **designation** (dimension short code)
- **description** (name that describes the dimension).

MEMBER

Members are specific characteristics of the dimension.

Each member is composed by:

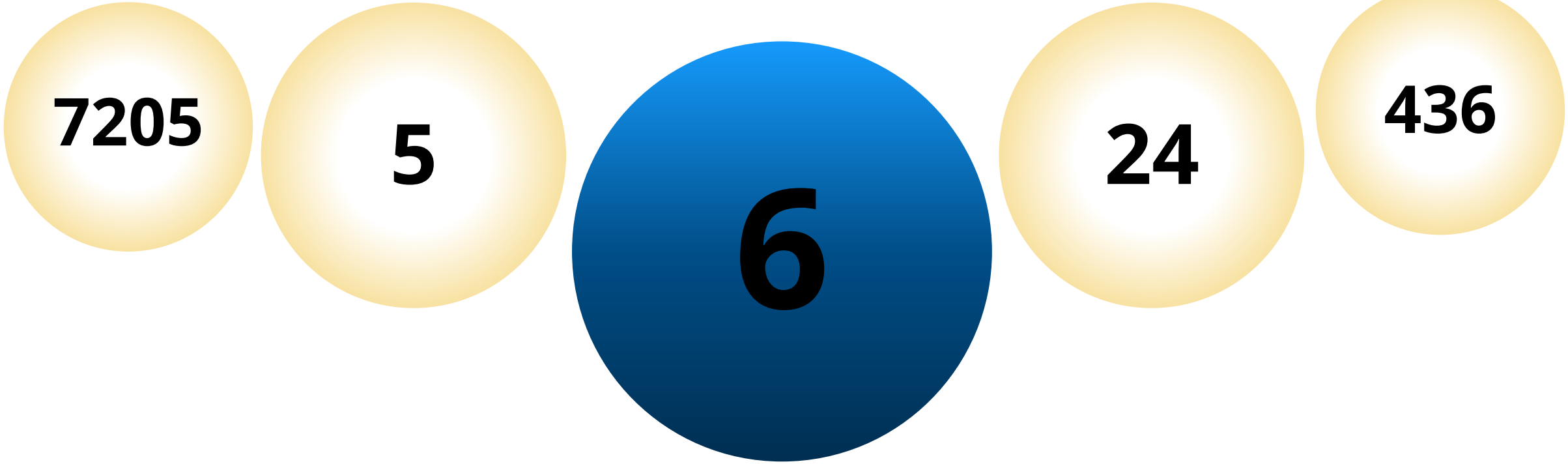
- **code** (member short code)
- **name** (member code description).

DEFINING A COMMON SET OF CONCEPTS

B) COLLECT STATISTICAL OUTPUTS

DEFINING A COMMON SET OF CONCEPTS

B) COLLECT STATISTICAL OUTPUTS

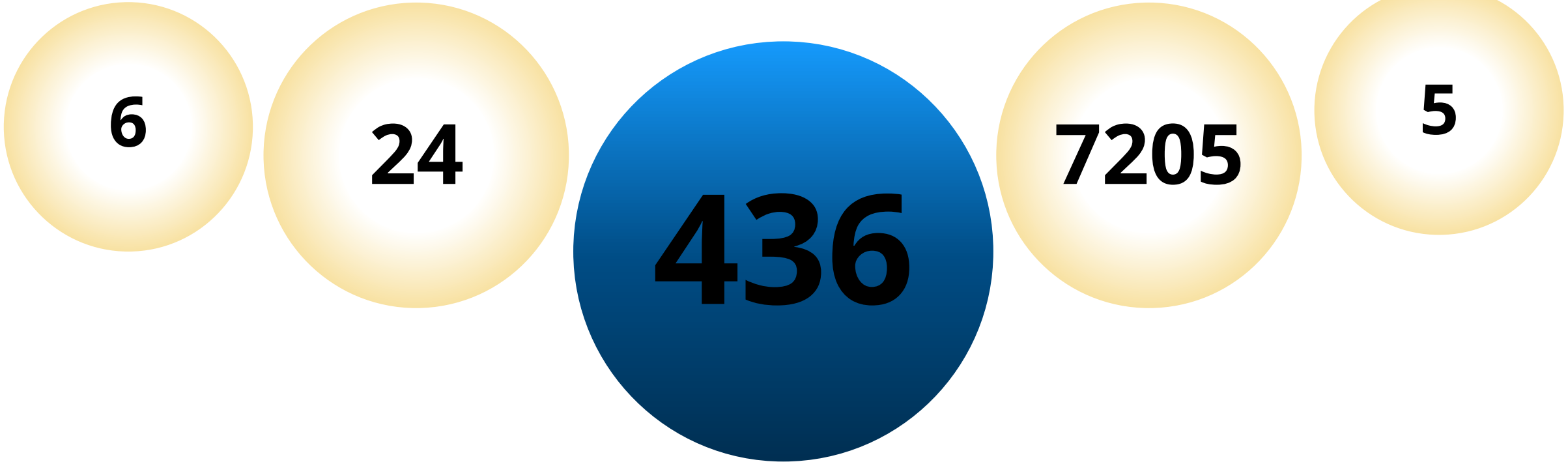


DEFINING A COMMON SET OF CONCEPTS

B) COLLECT STATISTICAL OUTPUTS

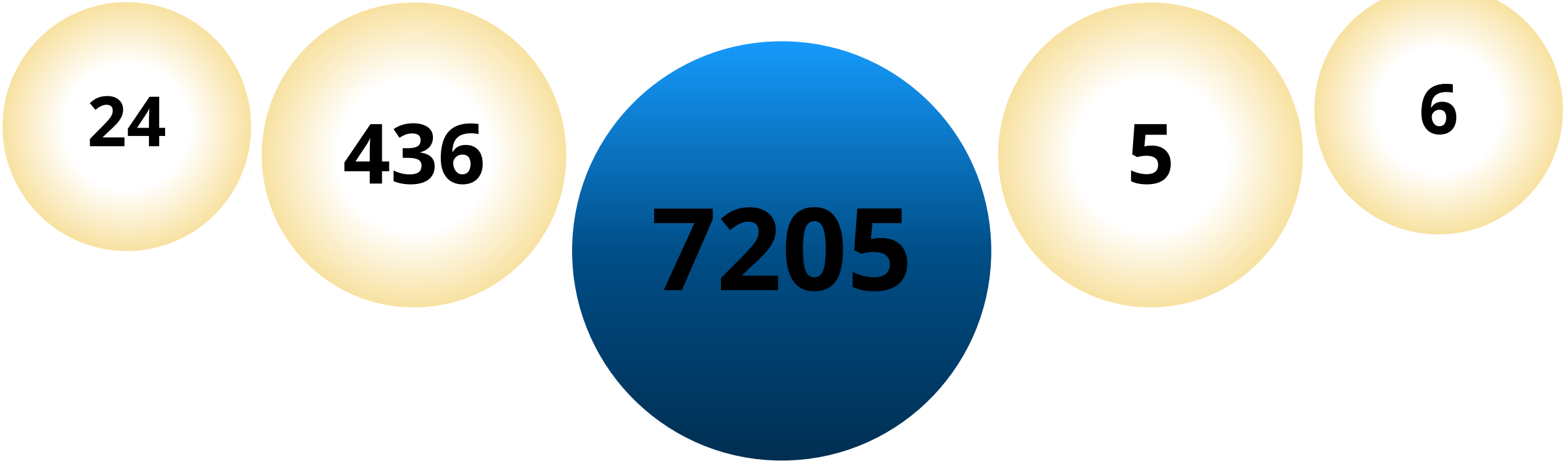
DEFINING A COMMON SET OF CONCEPTS

B) COLLECT STATISTICAL OUTPUTS



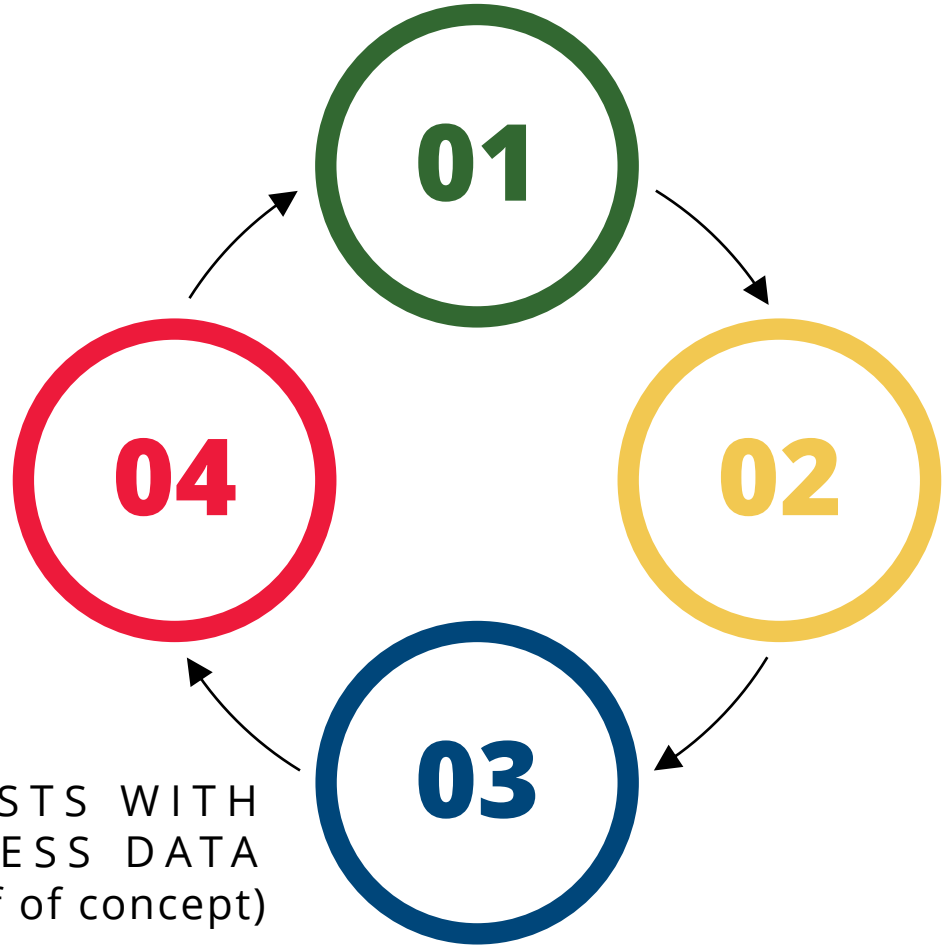
DEFINING A COMMON SET OF CONCEPTS

B) COLLECT STATISTICAL OUTPUTS



DEFINING A COMMON SET OF CONCEPTS

C) HARMONIZATION CYCLE



ANALYSE DIMENSIONS AND GROUP ALL MEMBERS

HARMONIZE MEMBERS AND STUDY CONCEPTS

- ✓ Define mapping tables
- ✓ Define aggregation tables

TESTS WITH BUSINESS DATA (Proof of concept)

INCORPORATE IMPROVEMENTS (RETEST THE RESULTS) AND DEFINE GOVERNANCE MODEL

DEFINING A COMMON SET OF CONCEPTS
D) RESULTS

DIMENSIONS

436

HARMONIZED DIMENSIONS

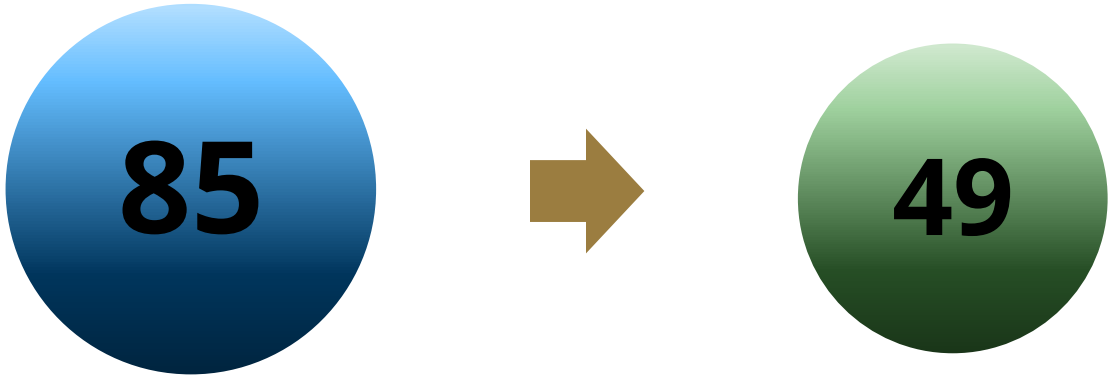
85

-81%

DIMENSIONS

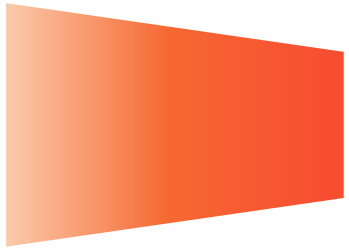
DEFINING A COMMON SET OF CONCEPTS

D) RESULTS



Example:

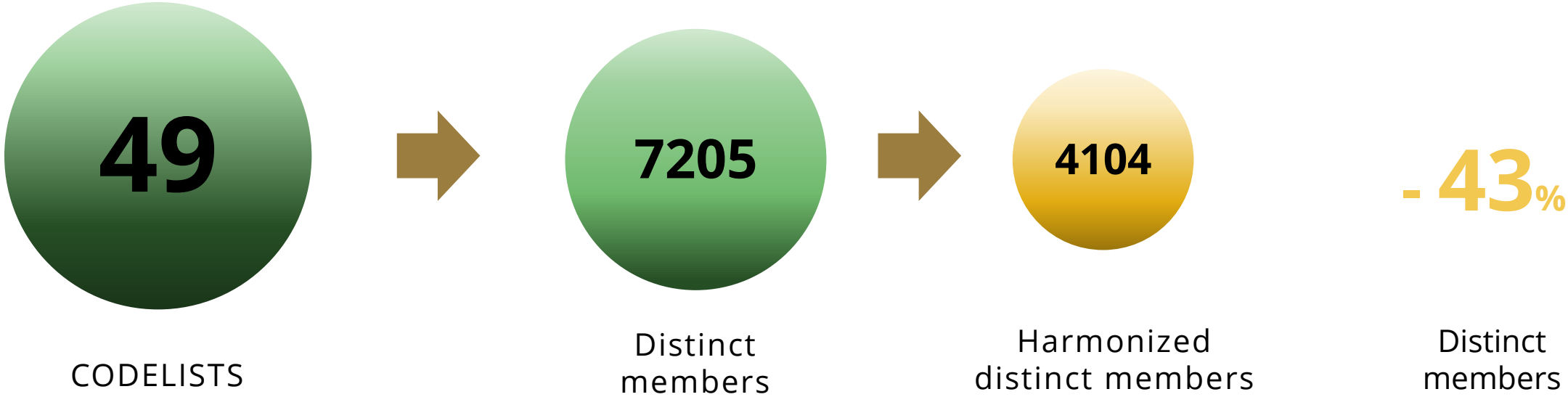
Original maturity
Residual maturity
Fixation rate maturity



Maturity

DEFINING A COMMON SET OF CONCEPTS

D) RESULTS



DEFINING A COMMON SET OF CONCEPTS
GOVERNANCE PRINCIPLES

USE ALPHANUMERIC CODES, STARTED WITH A LETTER.

FOLLOW THE STANDARD SDMX CODELISTS DEFINED IN THE SINGLE DATA DICTIONARY (SDD).

FOLLOW OTHER CODELISTS IN SDD, IF THERE IS NO MATCH FOR THE CONCEPT IN SDMX.

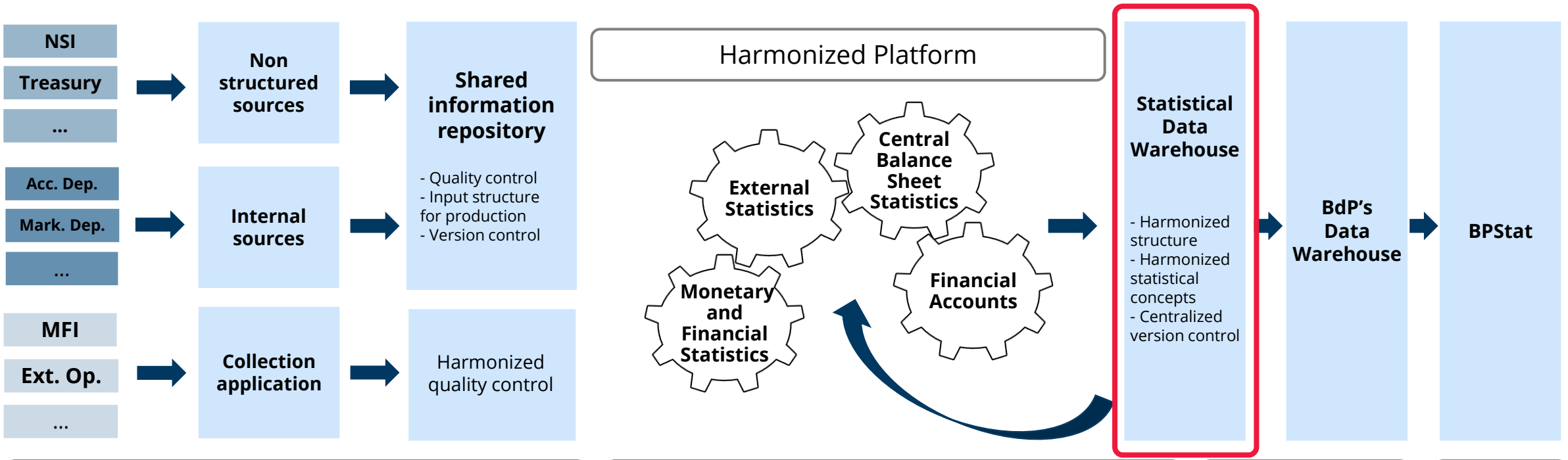
USE INTELLIGIBLE DESIGNATIONS AND CODES.

DESCRIBE THE UNDERLYING RULES FOR THE CODES OF EACH DIMENSION.

LIMIT CODE SIZE.

DEFINING A COMMON SET OF CONCEPTS
GOVERNANCE PRINCIPLES: WHERE TO START?

Reference information + Data catalogue



Simplify, Harmonize and Enable.

DEFINING A COMMON SET OF CONCEPTS

GOVERNANCE PRINCIPLES: WHERE TO START?

STATISTICAL DATA WAREHOUSE.

REFORMULATION OF STATISTICAL PRODUCTION PROCESSES.

PRODUCTION OF NEW STATISTICS.

NEW DIMENSION AND MEMBERS.

DEFINING A COMMON SET OF CONCEPTS
IMPLEMENTATION

WITHIN PRODUCTION PROCESSES

4 processes already adopted the new set of concepts

WHEN STORING STATISTICAL OUTPUTS IN THE DATAWAREHOUSE

3 statistical outputs translated into the new set of concepts

WORK IN PROGRESS

Planned integration of all statistical outputs translated into the new set of concepts by the end of 2024

CLOSING REMARKS

05

CLOSING REMARKS

SEMANTIC INTEGRATION IS A CORNERSTONE OF OUR STRATEGY.

CENTRALIZED STATISTICAL PRODUCTION PLATFORM.

CROSS-REFERENCE OF DIFFERENT DOMAINS (COHERENCE AND OVERALL QUALITY).

HARMONIZATION IS ALWAYS AN ON-GOING PROCESS.

IMPLEMENTATION IS STILL A CHALLENGE: A STEPWISE APPROACH WAS CONSIDERED MORE FEASIBLE.

“COMMON LANGUAGE”.


QUESTIONS