




42nd Central Banks Governors' Club Meeting of the Central Asia, Black Sea Region and Balkan Countries

Data management in the context of monetary policy and central banks' activities

Theodoros Mitrakos
Deputy Governor, Bank of Greece

Prague, 20-22 November 2019

- 
- A large, stylized black outline of an unrolled scroll, containing a list of four bullet points.
- Central bank activities and the role of statistics
 - Euro area statistical data projects
 - Data management at the Bank of Greece and selected case studies
 - Challenges – the way forward

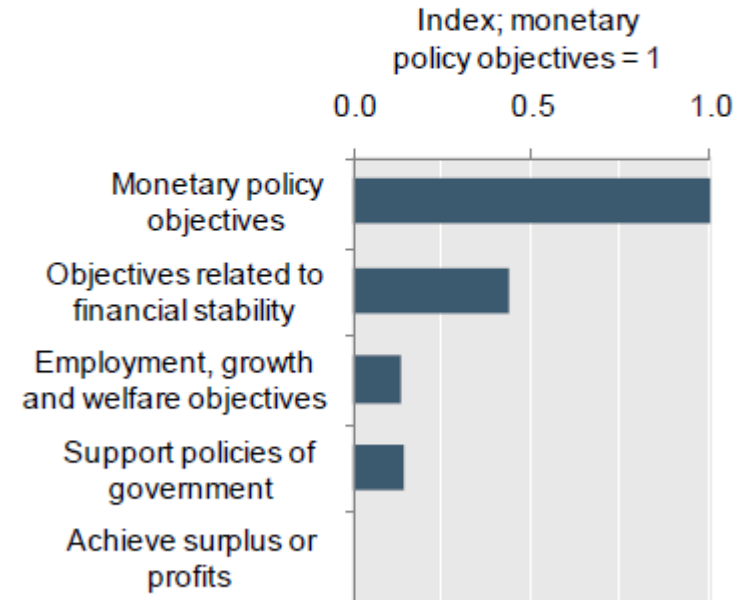
The mandate of central banks drives data needs and calls for the development of statistical databases



- Central banks have as a primary objective the maintenance of price stability, followed by financial stability. Other objectives could include:
 - Promoting employment and growth;
 - Support government policies;
 - Oversight of payment system infrastructure.
- Statistical data and related analysis represent a critical factor in most central bank functions:
 - Monetary policy implementation;
 - Banking and insurance supervision;
 - Macroprudential policy;
 - Payment systems oversight;
 - Lender of last resort function.
- Central banks are both a big user (e.g. national accounts) and a producer of statistical data series (e.g. monetary and financial institutions statistics, external statistics).

Weight of central bank objectives in central bank laws

Per cent of 47 central banks



Source: BIS analysis of central bank laws.

Data classification covers multiple dimensions interlinked with central bank activities



BANK OF GREECE
EUROSYSTEM

Granularity

- Aggregate data (e.g. GDP, total deposits)
- Data per reporting entity (e.g. bank)
- Micro-level data (e.g. AnaCredit)
- Survey data (e.g. household & consumption survey)
- Big data

Purpose

- Monetary policy
- Economic analysis & research
- Statistical purposes
- Prudential supervision (e.g. banking, insurance, etc.)
- Oversight of payment systems
- Collateral valuation
- Market conduct
- Consumer protection

Nature

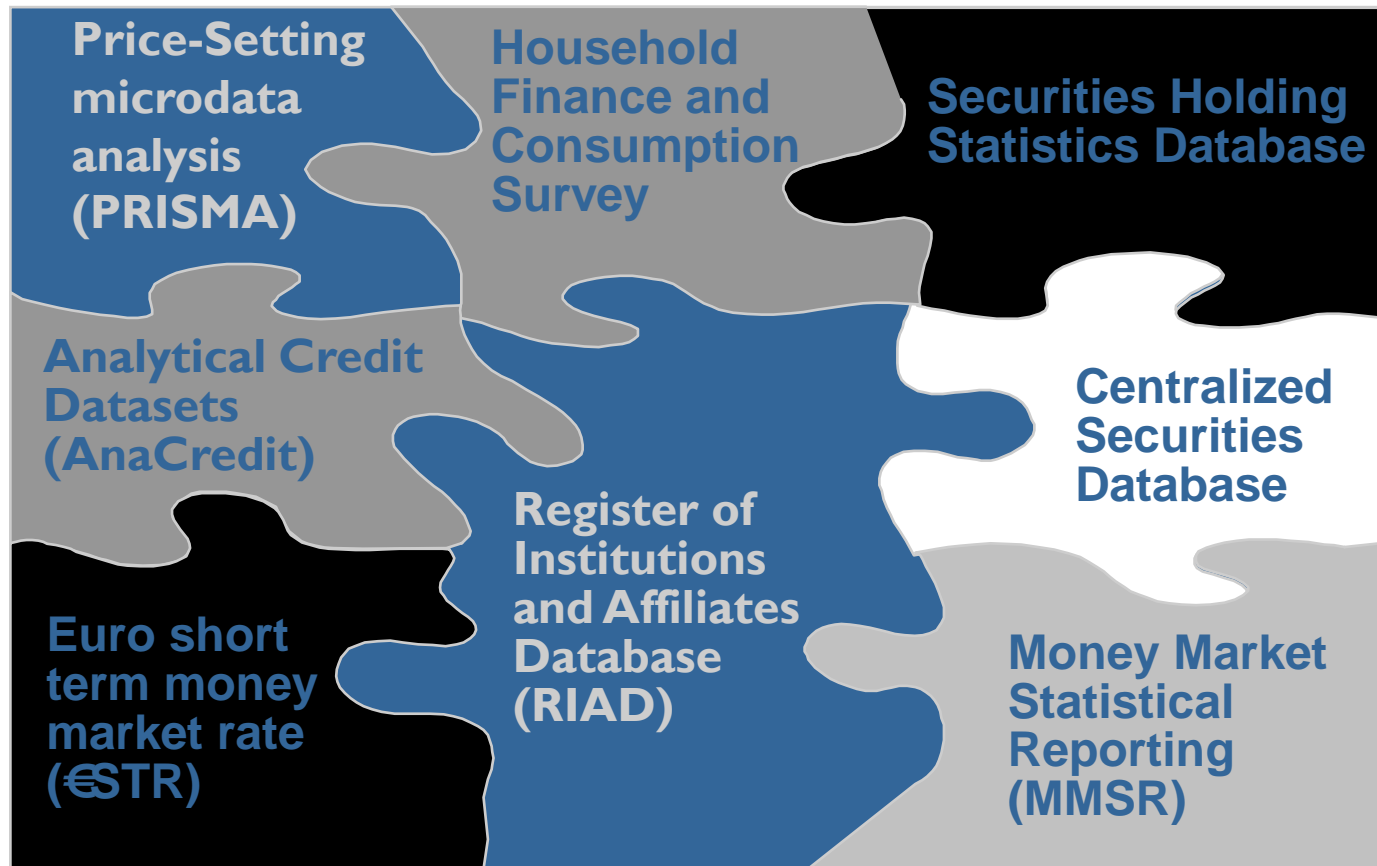
- National
 - Residential property data
 - NPL servicers data
- Euro-area
 - AnaCredit / RIAD
 - Securities Holdings Statistics Database
 - Price-setting microdata analysis
- Global
 - BIS statistics
 - IMF financial soundness indicators

The ESCB has many statistical projects to collect fully harmonized and standardized granular data



BANK OF GREECE
EUROSYSTEM

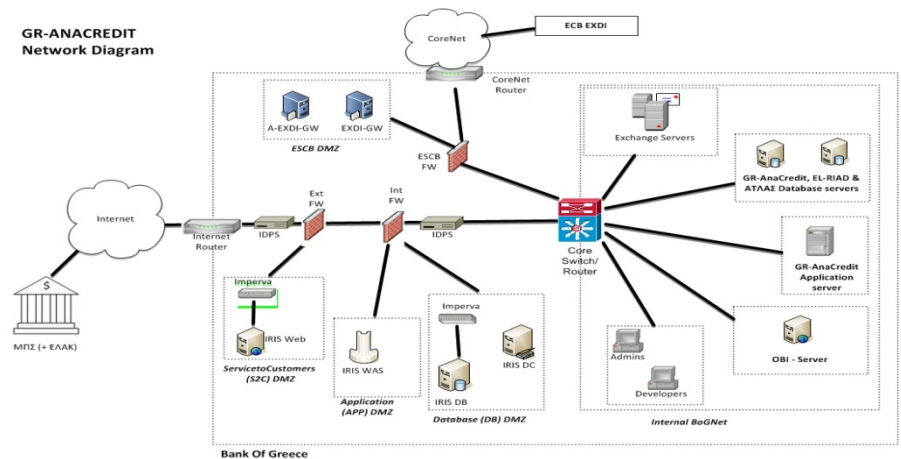
- Good quality and timely statistical data contribute to informed policy decisions.
- The global financial crisis triggered new policy questions; need to zoom in beneath aggregates.



AnaCredit is a key building block in the ESCB data strategy



- AnaCredit stands for analytical credit datasets and became operational as of Q4 2018.
- The AnaCredit database contains detailed, timely and harmonized loan-level information on bank loans of euro area credit institutions.
- AnaCredit contains credits with a minimum amount of €25 thousand extended only to legal entities.
- AnaCredit requires ca. 94 data “attributes” and 7 unique identifiers covering many aspects of the exposure:
 - Counterparty (e.g. LEI code, address)
 - Instrument (e.g. type, currency, NPL status)
 - Collateral (e.g. type of protection, location of collateral)
 - Accounting data (e.g. accumulated impairment).

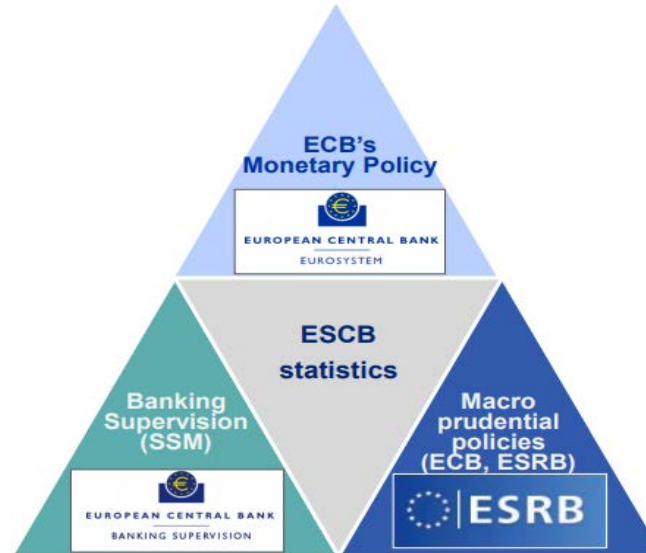


AnaCredit serves as a magnifying glass for several central bank tasks; a multi-purpose dataset



BANK OF GREECE
EUROSYSTEM

- AnaCredit supports many central bank functions:
 - Monetary policy analysis and implementation;
 - Financial stability;
 - Risk management;
 - Economic research;
 - Statistics.



- Harmonized data and a complete coverage for all euro area countries allows:
 - Assessment of developments in credit demand and supply (including SMEs);
 - Accurate exposure concentration analysis (sectoral, regional etc.);
 - Evaluation of total company exposure towards all euro area banks.
- *Non-standard monetary policy measures and the close link between monetary and macroprudential policies increase the importance of AnaCredit.*

RIAD facilitates integration of various datasets; other projects tailored to specific needs



BANK OF GREECE
EUROSYSTEM

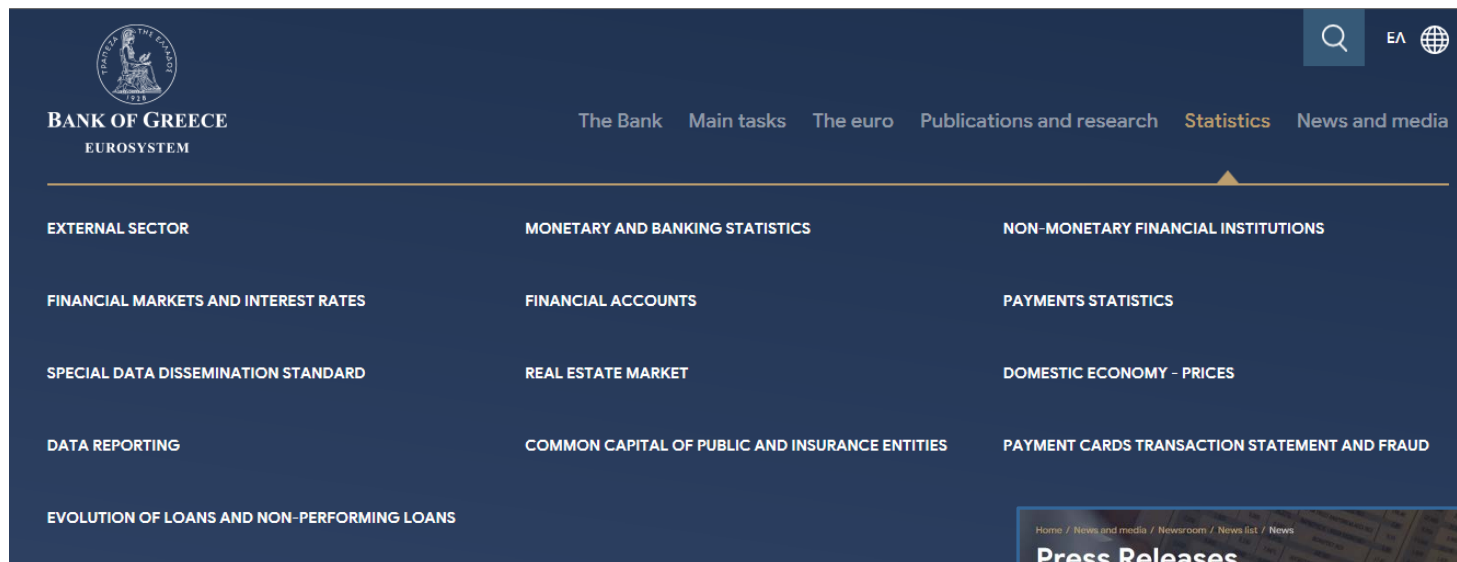
- The **Register of Institutions and Affiliates Database (RIAD)** is a business register, a shared dataset of reference data on legal entities, including non-financial corporates.
 - RIAD contains a wide range of attributes on individual entities and relationships between such entities that enables the derivation of group structures.
 - RIAD facilitates the integration of a variety of datasets, in particular by providing common identifiers.
- The **'Price-setting microdata analysis' (PRISMA)** network aims to collect and analyze the underlying price quote (micro) data from the harmonized index of consumer prices (HICP) and the producer price index (PPI) across the euro area.
- The **Household Finance and Consumption Survey (HFCS)** is a harmonized survey that collects microdata on the assets, liabilities, income and consumption of European households. The dataset provides insights into their economic behavior and financial situation.

Communication of statistical data is of paramount importance for central banks – the case of BoG



BANK OF GREECE
EUROSYSTEM

- The Bank of Greece collects, analyses and disseminates a plethora of data sets covering many areas; they feature prominently at the BoG website.

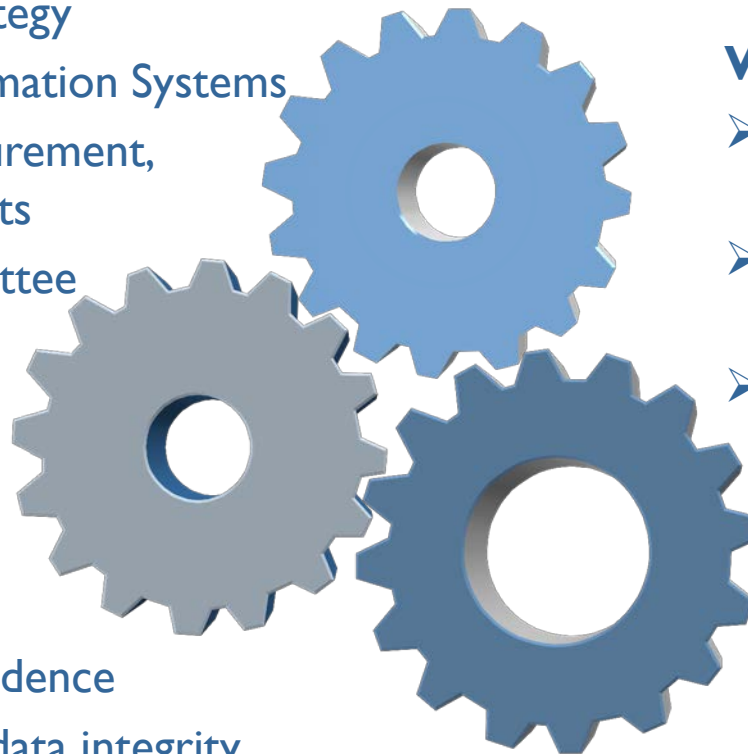


- The data communication strategy includes press releases, dedicated periodical publications, reference / analysis in core reports as well as working papers & research.
- Data collection, processing, analysis and dissemination are considered as a “public good”.



Multidisciplinary committees play a pivotal role in the design and implementation of data management strategy

- Committee for Information Systems
- Committee for Procurement, Contracts and Projects
- Business Plan Committee



Principles

- Professional independence
- Quality assurance / data integrity
- Confidentiality / access rights
- Timeliness & consistency
- Integration into decision making process

Well-defined roles

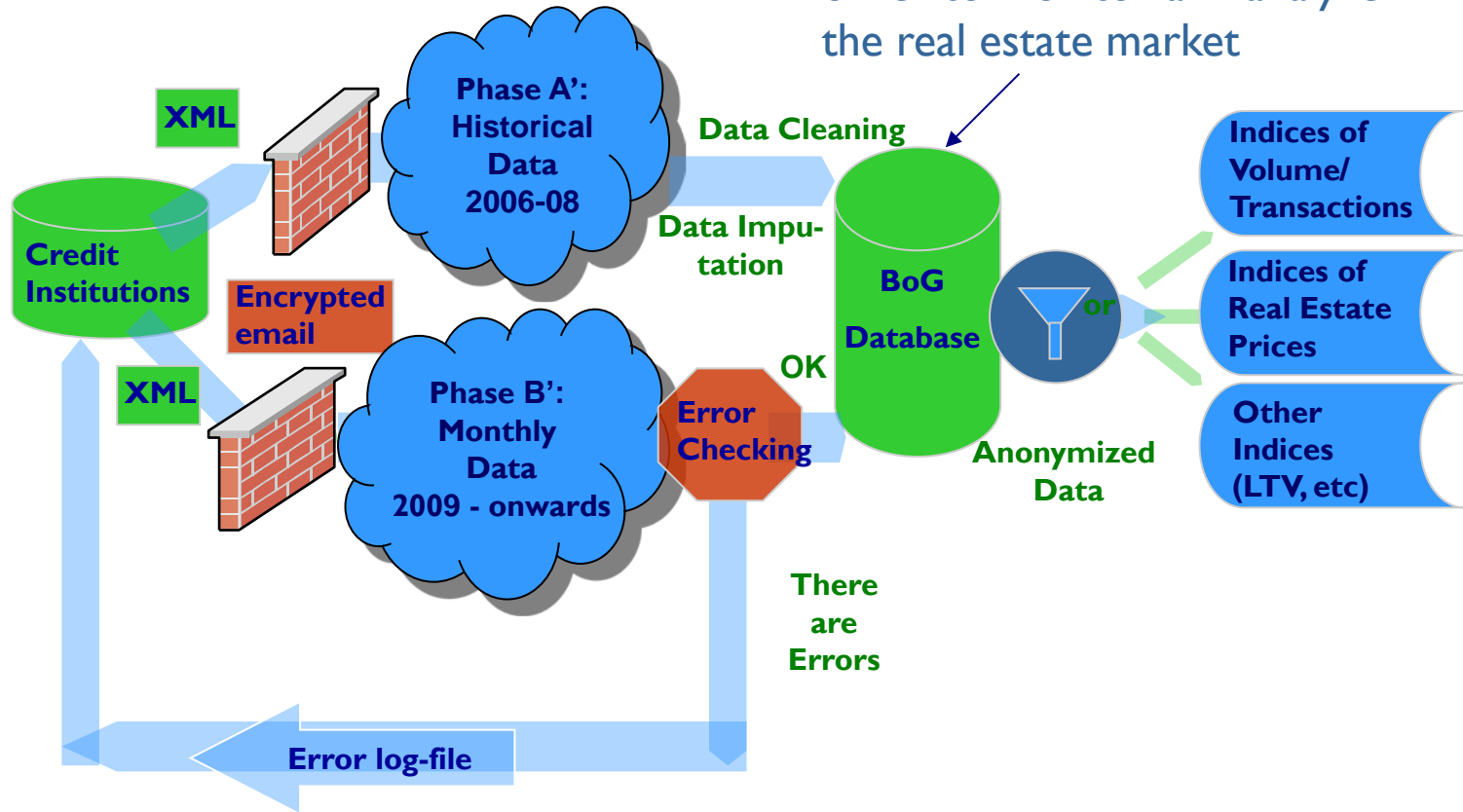
- Data owner (e.g. Statistics Dept. ; Supervision Dept.)
- System developer & owner (mainly IT Dept.)
- Special functions
 - Communications Dept.
 - Internal Audit Dept.

**Connection and
Authorized Access to
Statistical Data project
(COAST)**

Data management process: an indicative flowchart from Residential Property Data collection at the BoG



Bank of Greece Database in order to monitor and analyze the real estate market

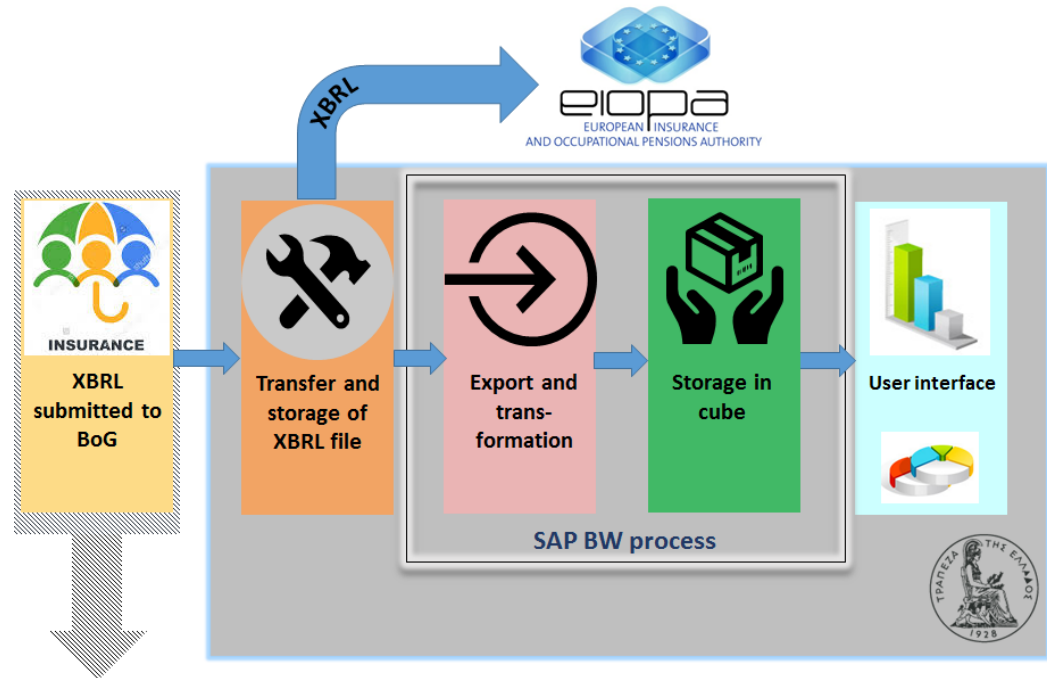


Integrated data management: the example of insurance corporations' statistics



- The advantage of granular datasets is that they can serve multiple users with the appropriate flexibility.
- The introduction of Solvency II entailed an enhanced and detailed reporting by insurance undertakings supervised by the Bank of Greece.
- The IT Department designed an integrated statistics and supervisory reporting for private insurance companies based on XBRL technology:

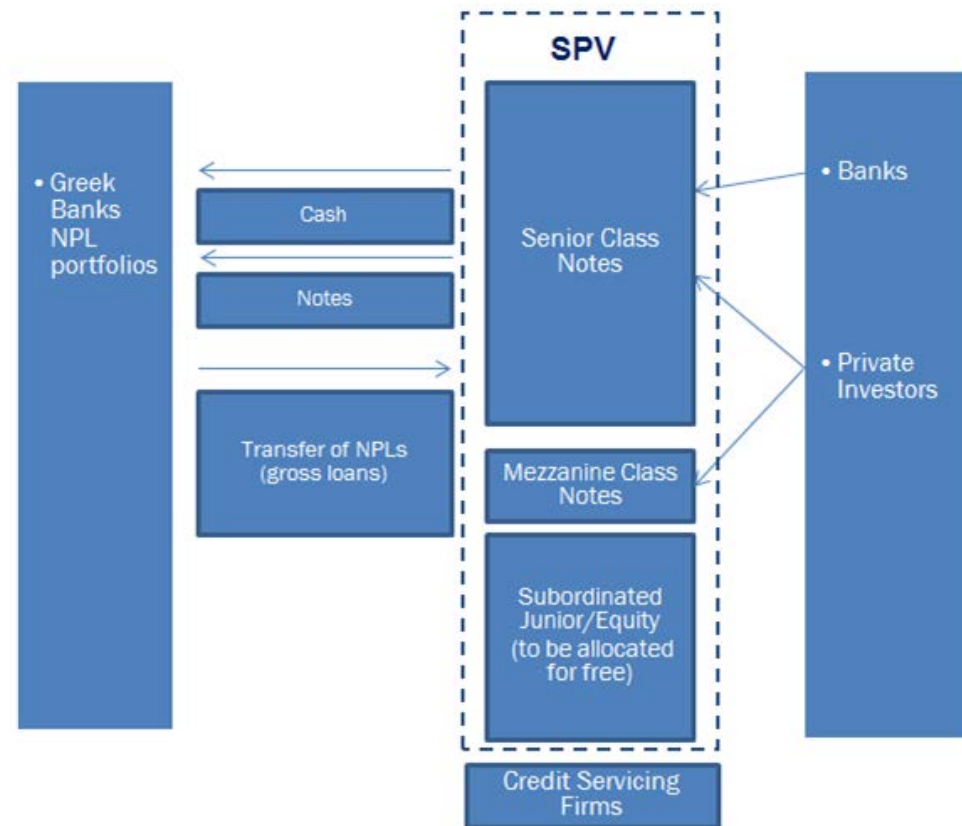
- Reduced supervisory burden for reporting institutions;
- High quality assurance and embedded support from IT;
- Timely and reliable data for all BoG functions & EIOPA/ECB.



Country-specific needs drive national initiatives

- **The elevated level of non-performing loans** remains the key challenge for the Greek banking sector and the Bank of Greece.
- Creating a secondary market for NPLs as a solution (securitization, outright sale, NPL servicing etc.)

➔ Credit Servicing Firms (CSFs) Data Project



- The Bank of Greece has decided to develop by 2023 an **In-house Credit Assessment System (ICAS)** to evaluate the financial health of corporations, in particular in the context of collateral management. AnaCredit and a **Centralized Balance Sheet Office** would provide necessary data input.

State-of-the art techniques: a robust machine learning approach for credit risk analysis



Irving Fisher Committee on
Central Bank Statistics

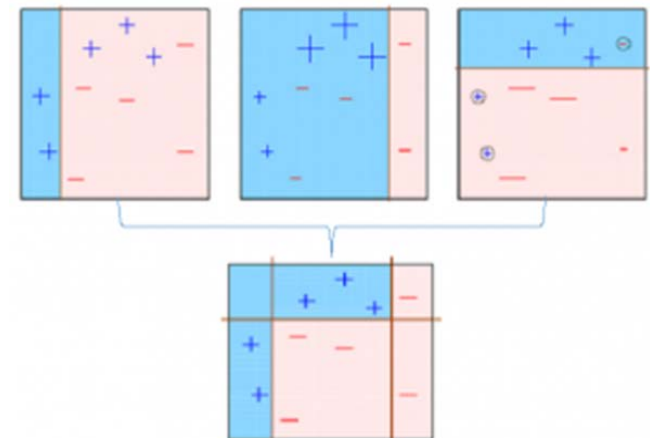
Ninth IFC Conference on "Are post-crisis statistical initiatives completed?"
Basel, 30-31 August 2018

A robust machine learning approach for credit risk analysis of large loan level datasets using deep learning and extreme gradient boosting¹

Anastasios Petropoulos, Vasilis Siakoulis,
Evangelos Stavroulakis and Aristotelis Klamargias,
Bank of Greece

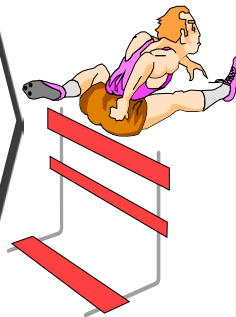
- Loan-level supervisory data of business loans above €1 million over a 10-year period are used to perform obligor credit quality classification and quantify Probability of Default.
 - A combination of data mining algorithms is employed with the aim to reduce dimensionality in the data and increase prediction accuracy.
-
- Key drivers include macroeconomic indicators and company financial ratios (ICAP database).
 - Cutting edge machine learning techniques like Extreme Gradient Boosting and Deep Learning Neural Networks outperform traditional Logistic Regression and Linear Discriminant Analysis.
 - Applications:
 - Development / challenge of IFRS/IRB models;
 - Leverage AnaCredit loan-level data.

Boosting: each model learns from the errors of the previous ones



Challenges

- Shift to granular/micro data
- Unstructured / high frequency data
- Market sensitive data
- IT infrastructure needs (systems & human resources).
- Growing importance of non-regulated players in the financial sector



Concerns / risks

- Model overfitting (out-of-sample failure)
- “Black Box” / identify causality
- Data confidentiality / access rights
- Overburden reporting institutions
- Validity of raw/web data

Holistic approach to data analysis

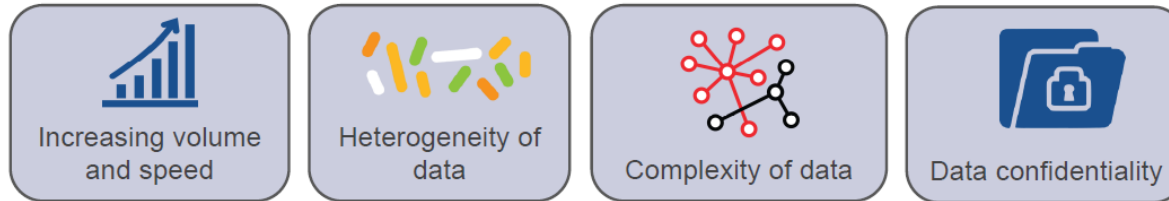
The way forward: new challenges call for a holistic approach in central banks data management



BANK OF GREECE
EUROSYSTEM

New challenges call for a holistic approach

To exploit the increasing availability of micro data challenges have to be addressed:



ECB data for analysis and decision-making

3

www.ecb.europa.eu ©