Economic and Financial Regulation in the Era of Big Data

Panel 1 : MEASURING

Renaud Lacroix

Director, Statistical and IT engineering division

24 November 2017
Salient features of BigData for measurement

- **New age of statistics:**
  - Need for more granular data to address new / complex issues
  - High expectations: ever-increasing volumes of data as input for better forecasts, indepth analysis in a more compressed time frame

- **Achieve multiple targets:**
  - **Handle heterogeneity**: BigData are aggregated from multiple sources
  - **Hedge against the uneven quality** of the data and noise accumulation: BigData contain various types of measurement errors, outliers and missing values
  - **Manage data complexity** and data dependence: selection bias, incidental endogeneity
  - **Break down data silos** and develop data sharing
  - **Reconsider traditional computing infrastructures**:
    - Need for new systems that support massively parallel data storage and processing
    - Need to ensure high levels of security and data confidentiality

- **We need a new system to meet these objectives**
The Datalake project

A multidisciplinary granular data platform supplying flexible and innovative services to Banque de France internal users in the first place

The platform provides key services in the area of data management and analysis to the benefit of all business lines of the Banque de France:

- Rethinking of data collection and storage to facilitate data processings and inference tools
- Development of automated data quality management for first-level quality checks
- Building of a common analytical platform covering a wide range of user needs
- Fostering synergies between directorates by sharing datasets beyond organisational silos

The project is on track for full implementation by end 2018

- A collaborative work promoted by DG-Statistics with the IT Directorate and all business lines involved
An overview of the Datalake ecosystem

REPORTING AGENTS: FINANCIAL INSTITUTIONS, NON FINANCIAL CORPORATIONS

The Banque de France Data Lake

- Internal BDF systems/Web sources
- BDF users
- IMF, BIS, Eurostat

Data collection

Feedback loop

Web Portal (OneGate)

Data Quality Management (first level)

Business application

Business application

Analytical Platform
(dashboards, dataviz, advanced tools)

Data Sharing Platform

Data Dissemination Portal

Data storage

Access for external researchers

Open Data Room

General public
Datalake: a technical view