

IFC Satellite Seminar on "Post-crisis data landscape: micro data for the macro world", co-organised with the Central Bank of Malaysia and the European Central Bank

16 August 2019, Kuala Lumpur, Malaysia

## Paradigms for modern statistical data frameworks and systems<sup>1</sup>

Michal Piechocki,

Eurofiling

---

<sup>1</sup> This presentation was prepared for the meeting. The views expressed are those of the author and do not necessarily reflect the views of the BIS, the IFC or the central banks and other institutions represented at the meeting.



**BIS IFC Satellite Seminar  
Post-crisis data landscape:  
micro data for the macro  
world**

16 August 2019 | Kuala Lumpur | Malaysia

**Paradigms for modern  
statistical data frameworks  
and systems**

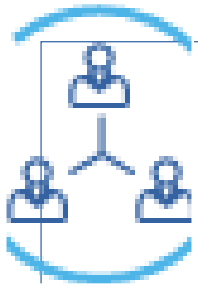
Michal Piechocki  
Board Member | Eurofiling Foundation p.f.  
Chairman | Frankfurt Group Technical Workshop



# Key study assumptions

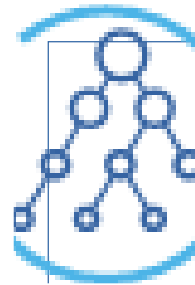
- Interdisciplinary design thinking
  - People, data, methodologies, standards, technologies, interfaces
- 6 qualitative case studies of international projects
  - ESMA TR, ECB SDD, EC FDS, NBS CRD, SCM NBFI, SG Banks
- Focus on 3 key contributing factors
- Practical framework of paradigms

# Three lenses



## Multi-stakeholder collaboration

- How stakeholders are organised?
- How they cooperate?
- How they remain engaged?



## Comprehensive data methodologies

- How to create cross-purpose definitions?
- How to reconcile legal and business differences?
- Use of logical description methodologies?



## Platforms and standards

- Technical interfaces and platforms?
- Use of international standards?
- Common identifiers?

# Micro data: gains and pains, the trade-offs

- Data as originally stored by FIs
  - By product, counterparty, transaction...
  - Cross-purpose analysis
  - Greater insight
  - Automated flagging power
  - Greater data sets
  - Raw data
  - Individual data
- Requires FIs input level definition
  - Variety of definitions of each
  - Purpose-driven regulation
  - Greater responsibility
  - False positives, false negatives
  - Scarce resources
  - Does not capture intentions
  - Personal data protection

# ESMA Trade Repositories Reporting



## Multi-stakeholder collaboration

- Reporting is performed by Central Counterparties and Trade Repositories
- Details of derivative transactions are reported to the Trade Repositories which are obliged to provide access to this data to the Competent Authorities



## Comprehensive data methodologies

- Tabular description of reporting obligations
- Each field to be reported described by proper data field and/or description on the ITS/RTS level
- Introduction of new legal frameworks based on EMIR allows for reconciliation of definitions on legal acts level



## Platforms and standards

- The reporting to TR may use different formats, such as ISO 20022, XML, CSV or FpML messages
- Single standard defined for direct transaction data feed to Competent Authority, as ISO 20022
- In case of EMIR reporting LEI is encouraged, with other identifiers accepted for specific cases only. For instrument reporting ISIN is endorsed

# ECB Statistical Data Dictionary



## Multi-stakeholder collaboration

- SDD activities performed by European Central Bank
- Introduction of complementary BIRD initiative allows for boarder cooperation within Expert Group, consisting of 10 National Central Banks and 30 Commercial Banks, coordinated by ECB
- The activities are further decided by BIRD Steering Group
- BIRD database is published and periodically updated on dedicated website



## Comprehensive data methodologies

- Creating precise definitions of the data through mapping multiple source dictionaries towards single unified dataset
- Each area is managed by one of dedicated ECB subgroups
- Current SDD consist of more than 14 000 elements used across multiple frameworks
- An example of granular framework incorporated into SDD is AnaCredit with nearly 130 unique fields (and 95 data attributes) used for description of detailed information on individual bank loans in the euro area



## Platforms and standards

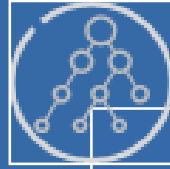
- Single Multidimensional Metadata Model (SMCube) to combine information usually modelled with multiple standards
- Describing multiple identifiers, with LEI and ISIN used as a common denominator for most frameworks

# EC Financial Data Standardisation



## Multi-stakeholder collaboration

- Supervisory Roundtables (ESMA, EBA, EIOPA, ECB, SRB)
- Lead by European Commission DG FISMA
- Subject-matter experts for data modelling
- Small group of experts for review



## Comprehensive data methodologies

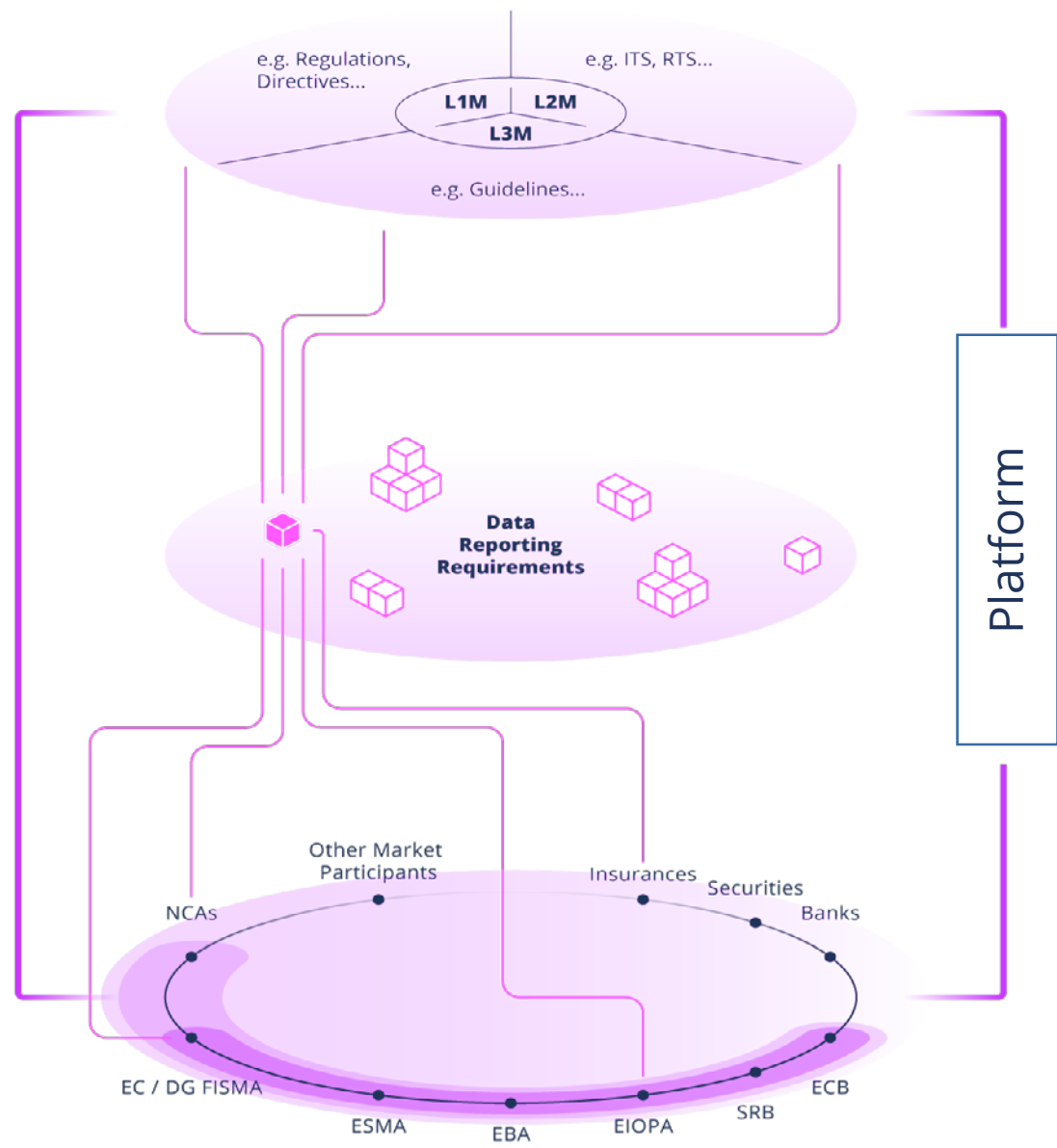
- Common definition of DRR
- One data description methodology for information requirements: Data Point Model
- 23 structured supervisory frameworks described
- Direct link to L1M, L2M, L3M legal acts and definitions



## Platforms and standards

- 9000+ DRRs
- 33 identifiers
- 7 reporting standards
- All models available on internal collaborative, metadata description platform used by subject matter experts



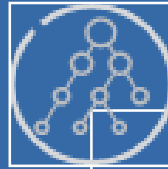


# NBS Credit Registry Dictionary



## Multi-stakeholder collaboration

- National Bank of Spain leads the effort
- The purpose: enable single standard for data collection
- Development started around credit registry dictionary
- Public consultation of taxonomies



## Comprehensive data methodologies

- Definitions started from loan-by-loan system and extended EBA (COREP/FINREP), BSI-MIR
- Single dictionary
- This allowed to extend into FINREP solo
- Methodology: Data Point Model
- Close tie to instructions / circulars (legal)
- Key data owners involved in definition of dictionary
- Top institutional leaders led the project



## Platforms and standards

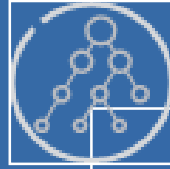
- Standards: SDMX, XBRL
- BdE developed an open source software

# SCM Private Retirement Schemes



Multi-stakeholder collaboration

- Fund management industries
- The purpose: standardised data collection
- Templates consulted with stakeholders
- Simple interface for data collection



Comprehensive data methodologies

- Methodology: Data Point Model
- Supervisor develops a single understanding of data requirements
- Open tables for threshold-restricted top transactions



Platforms and standards

- Standard of exchange: XBRL
- Best practices driven by experiences of the Bank of England
- Considerations of reuse of other regulatory dictionary for financial reporting

# Singaporean Banking Initiative



## Multi-stakeholder collaboration

- The purpose: to develop common understanding of supervisory and statistical MAS data requirements
- Lead by a group of private vendors and consultancies
- Founded by 8 FIs in 2016, 32 FI as of 2019
- FIs and Big4 comment and provide feedback on common draft models
- Regulator allowed to partially recover costs through grants



## Comprehensive data methodologies

- Methodology: Data Point Model
- Single data dictionary covering MAS 610 and 1003 requirements
- Reconciliation of differences through expert review
- Updated in iterations based on changes in legal requirements and system constraints
- Fis can extend the common dictionary to match their internal requirements



## Platforms and standards

- Key output: XBRL taxonomy
- Interoperable with SQL, JSON, XML, CSV
- Subject matter experts use common metadata modelling platform for design and review of data models

## Observations

- Five initiatives define requirements for micro data
- Five initiatives used a common open data design methodology: Data Point Model
- Four initiatives emphasized early engagement of relevant stakeholders in design phases of data requirements
- Two initiatives relied almost entirely on data requirements as defined by industry participants
- Two initiatives implemented collaborative platforms for metadata design

## **Observations (continued)**

- One initiative relied on a shared infrastructure for data collection
- Three initiatives relied on common international identifiers: LEI, ISIN (ISO standards)
- Five initiatives relied on open XBRL data collection standard
- Six initiatives facilitate linked business rules

# Framework of paradigms

- Collaboration with stakeholders
- Common data understanding
- Comprehensive data models
- Linked business rules
- Agility of data methodology
- Platforms supporting continuous feedback
- Technical interoperability
- Balance between policy and market practice



**Thank you**

---

**Contact**

Michal Piechocki  
[michal@eurofiling.info](mailto:michal@eurofiling.info)

