ECB Data Dictionary Roadmap

July 2015

The Authors

Arnold Veldhoen

Sergio Gianni

Jonathan Van Malleghem

The European Central Bank is continuing progress on its ambitious roadmap towards increasingly data-driven market monitoring, intervention and regulation. Under the umbrella of a Statistics work programme, a number of initiatives have been introduced which have a profound impact on those EU banks that are directly or indirectly subject to the SSM regime. This practice note summarises the ECB Statistics developments to date; the plans ahead for 2016-2018; and includes the views from key Reply practices in Europe. In particular, Italy is one of the most advanced countries in data-driven supervision with its Puma2 project. The ECB is looking at this as a best practice standard with potential wider rollout across the EU area. We strongly recommend all clients take note of the whole of the ECB 2016-2018 programme in order to align their respective Risk, Finance and Technology data architecture roadmaps with the increasing regulatory reporting requirements.

Background

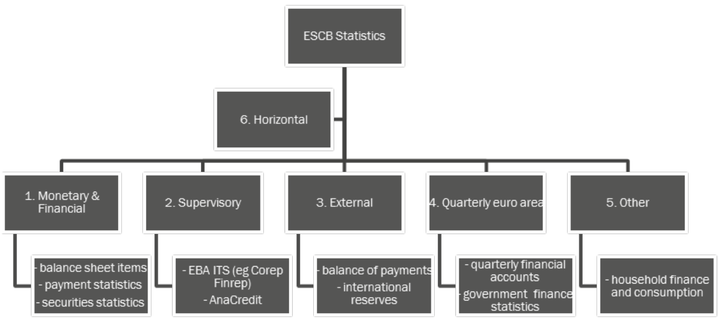
"The ECB has every interest in facilitating and promoting integration and standardisation, on the 'input side' too, within banks' internal systems, for only this will ensure coherent information."[1]

Mario Draghi

Introduction

The ECB has recently unveiled and reinforced a number of initiatives on its path to data-driven bank supervision. The drivers behind these initiatives and the practical impacts for countries and banks are outlined in this practice note.

The overall overview of the roadmap is in the 2015 ESCB Statistics programme[2] A stylised version of the programme follows:



ECB acknowledges that there are a large number of reports required, often with different publication or submission frequencies and different levels of aggregation. This leads to duplication and overlaps.[3]

A major point of contention for reporting banks lies in what ECB refers to as 'the use of micro-prudential data for macro-prudential purposes'. In practice, this means that SSM banks will have to submit increasingly granular loan level data so that ECB, for example, can aggregate that data for EU policies regarding economic growth of SME. This is one of the key drivers of the Capital Markets Union initiative.

[1] *Seventh ECB Statistics Conference, Frankfurt am Main, 15 October 2014. https://www.ecb.europa.eu/press/key/date/2014/html/sp141015.en.html*

[2] *2015 ESCB Statistics Work Programme*

[3] *European Reporting Framework (ERF): Key facts and information STC-SSM/2015/002.1*

REPLY
AVANTAGE

ECB Data Dictionary Roadmap

July 2015

The ECB programme encompasses a number of initiatives known as the ERF: the European Reporting Framework. The ERF needs to address the challenge of data gathering and submission by banks to their National Competent Authority (NCA); which in turn submits it to ECB, often using EBA validation rules. Different NCAs in the various countries are, however, at differing technological and functional maturity levels within their national infrastructures.

Input Layer and Data Dictionary

Under the catchphrase 'Input Layer', ECB is starting to develop more detailed data definitions. In the current reporting frameworks, banks use their own risk taxonomies because there is no internationally agreed or prescribed taxonomy. Regulations, Directives, Standards and Guidelines from the various regulators are often a combination of Word documents consisting of legal texts alongside Excel reporting templates. This often leads to a long list of Q&As before a consensus can be reached on how to interpret the texts and fill in the templates.

An important milestone in harmonising data models for over 120 banks has been the 2014 ECB Asset Quality Review data collection. This was, however, a one-off, highly manual, labour-intensive process. As an important next step, ECB is now embarking upon the AnaCredit process – a large data warehouse project that is set to collect four times the number of AQR data fields, and to do that consistently every quarter.

AnaCredit in a nutshell

Introduction

In line with the creation of a Capital Market Union ("CMU"), the ECB is leading the constitution of a central credit data register, referred to as Analytical Credit and Credit Risk Datasets ("AnaCredit"). The objective of this is to collect granular loan-by-loan credit risk information from Europe's financial institutions.

This will enable the ECB to analyse the institutions' data without relying on the ability of each body to respond to ad hoc requests. Even though the requirements are not yet finalised, the scope and timeline is taking shape. Institutions are already integrating this additional request into their data aggregation projects, such as BCBS 239 compliance programmes.

AnaCredit will require financial institutions to report fairly granular loan, facility and collateral information on a quarterly basis, in line with the Asset Quality Review exercise ("AQR") conducted in 2014. Acknowledging the difficulty of institutions to respond to the AQR in time and with sufficient data quality, the ECB will phase in the requirements. However, structural issues such as the dependence on manual workarounds or the capacity of institutions to report on granular loan-by-loan line levels represent a significant challenge for the coming years.

The timeline considered and the requirements for each phase are highlighted in the figure below:

	January 2018 *Q4 2017 data*	End 2019 *Q2 2019 data*	October 2020 *Q3 2020 data*
Threshold	+25K by debtor	+25K by debtor Branches outside EU: +1Mlo EUR by debitor	No exemptions +25K by debtor Branches outside EU: +1Mlo EUR by debitor
Format	Not Borrower by borrower but **Loan by Loan**	Loan by Loan	Loan by Loan
Scope	Domestic + branches Legal entities Deposits, loans and commitments Individual reporting	Domestic, branches and subsidiaries Legal entities Loans, Off Balance and credit derivatives Consolidated reporting	Domestic, branches and subsidiaries **All household loans** and legal entities All financial instruments Consolidated reporting
Submission	Q+40 WD M+30 WD – Other data	To be assessed	To be assessed

2

Figure 1 – AnaCredit timelines

The AnaCredit project, however, typically only covers the Credit Risk aspect of a bank risk taxonomy. Market Risk and Operational Risk require similar data standards: we only have to consider, for example, the reporting needs resulting from the Fundamental Review of the Trading Book for Market Risk or the new Standard Approach for Operational Risk, both part of the Basel Capital Floors programme. Under the ECB STE programme and the reports required by EBA SREP standards, those risk types are also increasingly prescribed on a low level of detail and the industry is becoming more and more familiar with the ECB STE templates. For example on LCR, Funding Plans or Operational Risk. Templates largely stem from EBA consultations.

The next step that ECB plans to take is the Banking Data Dictionary initiative. In April 2015, it held a two-day workshop on the topic with selected countries and central banks. There was a particular focus on the systems already in place in Austria and Italy. The Data Dictionary model looks primarily at data structures used in Corep, Finrep, AnaCredit, Money Market Statistics and Securities Holdings.

A number of complexities can be flagged up as we move towards the standard ECB Data Model:

- Each bank has individual ways to identify clients and products, although projects like FSB reporting-driven LEI (Legal Entity Identifier) or UTI (Unique Transaction Identifier) will increase standardisation over time;
- Reporting banks often have a number of systems and data models in use themselves. We do note the increasing integration of Finance and Risk data management at banks, in particular to comply with expectations such as BCBS239 or as an opportunity to face the challenges of the Digital Banking revolution;
- Consolidation requirements differ between Regulatory and Accounting reporting sets and purposes. Intercompany exposures are needed on sub/ branch level, particularly in reports used for Recovery/ Resolution data gathering or Liquidity. Such exposures are eliminated in consolidated reports such as Large Exposures. Changes in, for example, IFRS rules over time or time-series disturbing bank acquisitions and divestments are just too complex to be easily integrated into a universal data model. It is worth noting, for instance, that the European IFRS9 endorsement expected later in 2015 will mean a fundamental shift in Credit Risk models and data with regards to provisioning.

Interestingly, the Data Dictionary project for the time being distances itself from prescribing technical standards such as XML or XBRL in the same way as we have already seen in other EU standards (SEPA is a good example). We do expect, however, that such standards will become inevitable in order to operationalise the framework. The main standard mentioned to date is the SM Cube methodology using VTL syntax, taking into account ISO20022 and BCBS239.[4]

Spotlight on Italy

In Italy, the financial, statistical, prudential and accounting reporting of supervised Italian institutions is founded on an integrated process called PUMA2. Structure, development and maintenance are coordinated by the Bank of Italy. The overall purpose of PUMA2 is to achieve standardisation, uniformity and regulatory compliance.

A panel of Banks' members of the PUMA2 Interbanking Group has been established. Members of the panel collectively determine the scope of the functional and technical requirements. While PUMA2 has been operational for more than two decades, the flexible use of defined rule tables (c. 100) and common software (c. 30,000 routines) have allowed PUMA2 member banks as a whole to adopt local and EU regulatory and supervisory specifications relatively quickly.

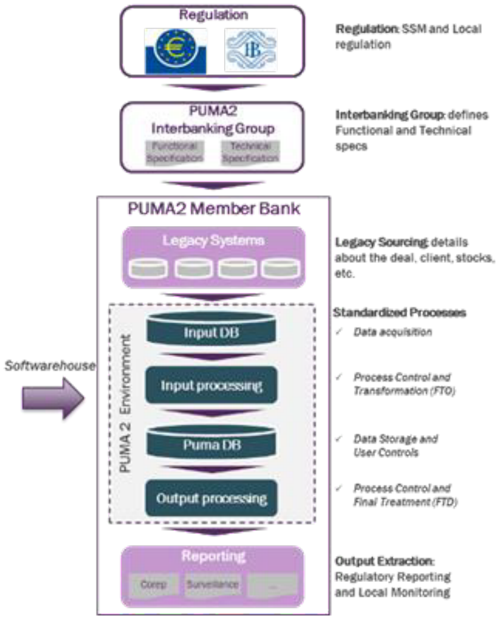


Figure 2 – Simplified process view of PUMA2

In addition to standardising the reporting process, the PUMA2 concept has allowed Italy to achieve what can be considered a de facto first version of a fully regulated Data Dictionary. In fact, PUMA2 uses procedures that today cover approximately 2,000 defined data fields.

Looking forward, work is currently being carried out by the Interbanking Group to allow PUMA2 to integrate new specifications and SSM enhancements. Overall, PUMA2 offers an interesting example to regulators and supervisors as they debate how to approach the ECB Data Dictionary.

[4] *Towards a Banking Data Dictionary, EBF_015072*

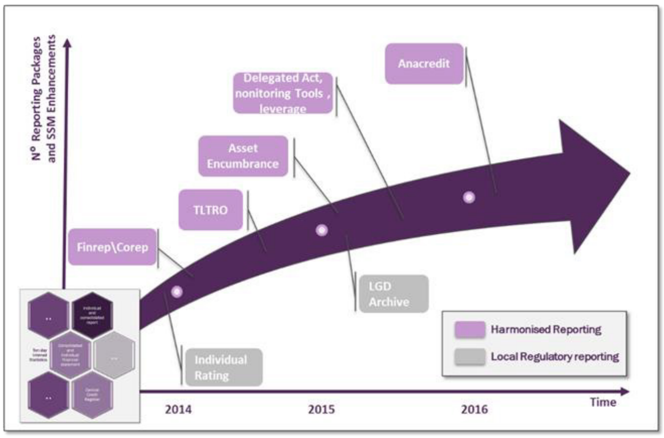


Figure 3 – Evolutionary trajectory of PUMA2

The German landscape

The Bundesbank called a working group meeting in May this year on AnaCredit and the ECB Data Dictionary. In alignment with the ECB roadmap, the Bundesbank has an additional national transition path where the objective is to implement some of the ECB requirements earlier on a national level. In particular, data for private real estate financing and data for private households are to be implemented in Germany before 2020 (Phase III of ECB roadmap).

The technical concept for Bundesbank reporting for AnaCredit is largely the same as the existing one for the national "Million Loan" reporting. This latter is on an xml basis, with automated data submission to Bundesbank. In the first phase, this national loan reporting is parallel to the AnaCredit data submissions. The long-term Bundesbank objective is the data integration of statistical and supervisory data as discussed at ECB level.

For regulatory reporting in Germany, most institutions use one of two standard regulatory reporting software products for the production of reports. Banks are fitting the data asks into the required data structure of either one of these two formats. Both are producing the XML or XBLR format for submitting to Bundesbank, which then submits them on to ECB.

Other Country Registers

Not all countries are at the same level of maturity in regards to national central credit registers or harmonised central bank reporting. In Spain, for example, the CIRBE is reasonably well developed; whereas in The Netherlands, the BKR register is not used for all types of counterparties. A number of EU country central bank reports currently remain in national legacy formats such as XML rather than EBA XBRL.

What about that other ECB data warehouse?

Following the demise of the asset-backed securities markets in the sub-prime crisis, in April 2011 the ECB issued a letter of intent indicating its objective to make use of a new data warehouse. The ECB chose to become a client of the data warehouse on a non-exclusive basis once ABS loan-level data requirements became an ECB collateral eligibility criterion, provided that the data warehouse would comply with certain technical specifications.[5] In 2012, the European DataWarehouse GmbH was launched to provide the data-handling infrastructure for the ABS market. It became operational in January 2013 with the commencement of the loan level reporting requirements for RMBS and SME ABS. The market itself is, however, still plagued by a number of structural regulatory issues currently under reform by EBA and Basel (Simple Transparent Comparable) and by the European Commission (Capital Markets Union).

Global regulator alignment

Not only is ECB on a path towards granular data across the Eurozone, but a number of regulations around EMIR and Mifid II in the EU and Dodd-Frank in the USA also require increasingly granular reporting of Securities transactions and holdings - or Assets in general, as Mifid II also comprises commodities. The industry has repeatedly requested ECB and ESMA to align data models to reduce the burden for institutions. One particular example is the ECB Money Markets Statistics reporting, which is arguably similar to the ESMA EMIR demands. For now, the ECB Statistics programme does not seem to be technically aligned to other regulators' data standards, other than FSB's LEI programme. US-based banks particularly would benefit greatly from, for example, CCAR alignment.

We also note that the Single Resolution Board (SRB) is starting its operations and that various National Competent Authorities have started to define data definitions for Recovery/ Resolution reporting. A number of new data fields that were not present in Corep/ Finrep or Liquidity are now being requested; and the SRB related reports need to be on a granular entity level – deconsolidating and including intergroup transactions and liquidity. We expect that ECB and SRB will align data models over time.

[5] *https://www.ecb.europa.eu/paym/coll/loanlevel/data_inf/html/index.en.html*

Impacts for our clients

We note that the ECB plans for a Data Dictionary will further drive data-driven supervision in Europe. The National Competent Authorities that roll out these requirements in national reporting systems will also demand additional reporting from banks not subject to the SSM regime. The Single Resolution Board data demands are expected to grow significantly during 2015 and 2016. Next year is also the year that BCBS239 will take effect, so banks are being required to increasingly integrate Finance and Risk data processes. Banks must continue to develop internal best practice risk taxonomies and internal data models now, before they are demanded by regulation.

One of the key risks we see is that banks may underestimate the size and complexity of the ECB data needs. Alternatively, they may be unable to respond in time due to legacy infrastructures and the amount of manual intervention needed by expert teams before data can be provided to regulators. Banks that have already started to invest in strategic data management capabilities have a head start on those who have not done this sufficiently so far. Leaders will gain competitive advantages in the increasing size and the ‘evening out’ of the level playing field.

In any upcoming ECB Model Review, the ability to demonstrate sound data management of the models’ underlyings will be a crucial differentiator in determining the amount of Model Risk capital. Cleansing the data manually before serving as model input will be an increasingly less accepted practice.

Contacts

Avantage Reply (Belgium)

5, rue du Congrès/Congresstraat
1000 Brussels
Belgium
Tel: +32 (0) 2 88 00 32 0
E-mails: avantage@reply.com

Avantage Reply (Netherlands)

The Atrium | Strawinskylaan 3051
1077 ZX Amsterdam
Netherlands
Tel: +31 (0) 20 301 2123
E-mail: avantage@reply.com

Avantage Reply (France)

5, rue des Colonnes
75002 Paris
France
Tel: 33 (0) 1 71 24 12 25
E-mail: avantage@reply.com

Avantage Reply (United Kingdom)

38 Grosvenor Gardens
London SW1W 0EB
United Kingdom
Tel: +44 (0) 207 730 6000
E-mail: avantage@reply.com

Avantage Reply (Luxembourg)

46a, avenue J.F. Kennedy
1855 Luxembourg
Luxembourg
Tel: +352 26 00 52 64
E-mail: avantage@reply.com

Avantage Reply (Turin)

Via Cardinale Massaia, 83
10147 Torino
Italy
Tel: +39 011 29101
E-mail: avantage@reply.it

Avantage Reply (Milan)

Via Castellanza, 11
20151 Milano
Italy
Tel: +39 02 535761
E-mail: avantage@reply.it

Avantage Reply (Rome)

V.le Regina Margherita, 8
00198 Roma
Italy
Tel: +39 06 844341
E-mail: avantage@reply.it

Xuccess Reply (Berlin)

Mauerstrasse 79
10117 Berlin
Germany
Tel: +49 (30) 443 232-80
E-mails: xuccess@reply.de

Xuccess Reply (Frankfurt)

Hahnstrasse 68-70
60528 Frankfurt am Main
Germany
Tel: +49 (0) 69 669 643-25
E-mail: xuccess@reply.de

Xuccess Reply (Munich)

Arnulfstrasse 27
80335 München
Germany
Tel: +49 (0) 89 - 411142-0
E-mail: xuccess@reply.de

Xuccess Reply (Hamburg)

Brook 1
20457 Hamburg
Germany
Tel: +49 (40) 890 0988-0
E-mail: xuccess@reply.de