IFC-Bank of Italy Workshop on "Data Science in Central Banking: Applications and tools"

14-17 February 2022

Defining business transformation rules in a standardised format – a practical case[1]

Daniela Arru and Giulia Oddone,
European Central Bank

[1] This contribution was prepared for the workshop. The views expressed are those of the authors and do not necessarily reflect the views of the Bank of Italy, the BIS, the IFC or the other central banks and institutions represented at the event.

Defining business transformation rules in a standardised format: a practical case[1]

Arru Daniela, Oddone Giulia

Abstract

When talking about developing a data product at European level and collaboratively, the need and benefit of standardisation and harmonisation is often implicitly assumed, but this is not necessarily enough for a successful product. An important decision to be taken is the format that accompanies the standard, as the data product should combine business and technical aspects to gain a wider adoption and contribution. In this paper we explain why a standardised format with business characteristics is needed when it comes to regulatory reporting. The documentation of reporting instructions plays an important role in improving data reporting processes for financial institutions. Transformation rules are part of it and have been used as a way of formalising the regulatory instructions. Examples of documentation of transformation rules in Europe show that language and formats adopted often vary. In this paper we draw a set of principles for a standardised business-friendly format that bridges the needs of technical and business users. Such format should on one hand improve the ability of business experts to contribute and consume business transformation rules, and on the other hand support implementation of innovative solutions by technical experts.

Keywords: Banks' Integrated Reporting Dictionary (BIRD), transformation rules, Validation and Transformation Language (VTL), Extract Transform Load (ETL), business process, Business Process Model and Notation (BPMN), Integrated Reporting Framework (IReF)

JEL classification: G21, E50, C81, C82, C88, L17

Contents

[1]The views expressed are solely those of the authors and do not necessarily reflect the opinion of the European Central Bank.

1. Introduction

Over the past decade and in response to the financial crisis, authorities have increased their demand of data for statistical and regulatory purposes. Data requirements are often based on different definitions within and across regulations. On the side of financial institutions, the lack of standardisation leads to a considerable effort to design and maintain the data reporting process.

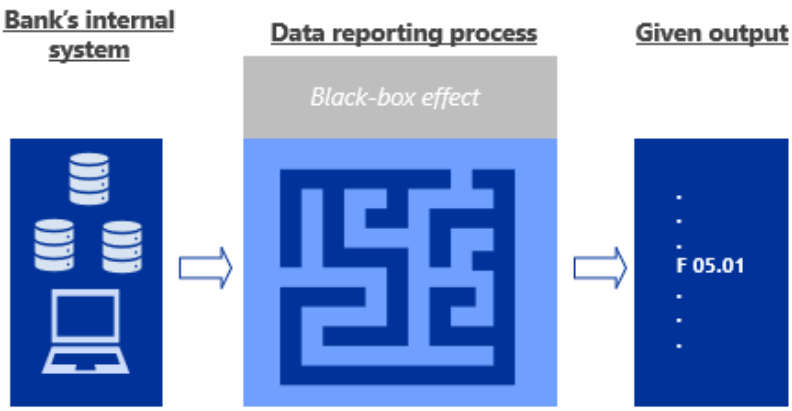
Every year each financial institution in Europe invests resources to interpret regulations and all associated manuals and instructions to fulfil their reporting obligations[2]. Financial institutions are often supported by software vendors and are dependent on proprietary IT solutions. Software vendors that compete to sell user-friendly and advanced analytical reporting solutions to reporting agents not only engage in software development, but also analyse and interpret once again and in their own ways the same regulations, manuals, and other material. Currently most parties involved in the business of regulatory reporting develop transformation rules independently. In this scenario of multiplication of effort on non-competitive tasks, transformation rules developed collaboratively and made available to the public have the potential to unfold competition and innovation.

Their importance stems from the fact that well-defined transformation rules provide semantically consistent interpretations of regulations, therefore increasing compliance with the regulatory requirements, reducing reporting burden, permitting data and attribute lineage, and improving the consistency and quality of the data reported to authorities.

In this paper we focus on the development of transformation rules given our experience in collaborative initiatives, and we explain why we consider the format as an important factor when we aim at increasing standardisation and at fostering innovation. Formats that are too technical, e.g. programming languages, impose barriers to the collaboration by business experts and have long developing and testing cycles. The adoption of business language, logical constructs and automatable and test-driven approaches ensures the right support to business experts. It also forms a sound and standardised basis for any technical implementation. A suitable format for transformation rules provides a publicly available standard and overcome the perception that transformation rules are a costly black-box.

Figure 1. Black-box effect in data reporting process

Simplified representation



[2] European Banking Federation (2019)

2. The role of transformation rules

Currently in Europe reporting agents submit on a regular basis information of their non-trading loans and advances. One example of such reporting is the template F 05.01 for financial information reporting of the European Banking Authority. The task of producing F 05.01 implies potentially thousands of different implementations of the data reporting process. Each implementation has different features, e.g. technical language, data dictionary and data model, ETL process design, data lineage. These features are the results of multiple factors, such as independent interpretations of concepts in the regulatory reporting documentation, banks' legacy systems and expertise developed by proprietary software solutions. A multitude of different implementations also means that switching from one solution to another can be extremely costly for financial institutions due to customisation of the reporting solution and dependencies with their internal ETL processes and data systems.

In some European countries[3] and at international level[4] collaborative initiatives have been introduced where experts from financial institutions and authorities analyse regulations, manuals and reporting instructions, apply their expert judgment and make their interpretation available for other parties to make use of it. They don't only find synergies on subject-matter analyses of highly complex regulatory documentation, they also provide an agreed, acknowledged and understood set of information to national market participants which in turn use it as their standard.

As an example, the Banks' Integrated Reporting Dictionary (BIRD) is a project established as a collaborative initiative between authorities and the banking industry in the field of regulatory reporting. BIRD develops a database encompassing a harmonised data dictionary, and a harmonised data model, that specifies how data can be extracted from the banks' internal IT systems to generate the reports required by authorities. BIRD is not limited to the description of data that banks need to collect from their internal systems. Experts also document in the form of transformation rules the operations to be applied to input data to enrich them with derived concepts so to have the complete set of data requested by authorities.

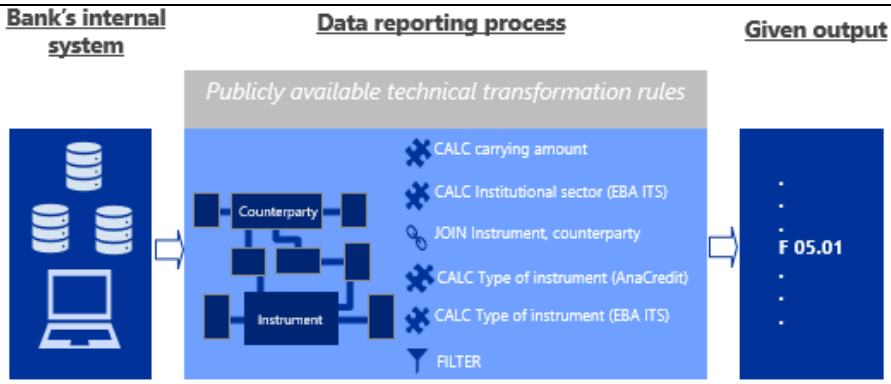
When thinking about template F 05.01, currently the process to create it can widely differ from bank to bank. BIRD provides one option of a harmonised data dictionary, including data model and transformation rules, for creating F 05.01 from a redundancy-free input. For example, in the BIRD database one can find calculations for the derivation of carrying amount, or for the institutional sector breakdown and the type of instrument according to European Banking Authority's reporting requirements, which may differ from the ones required for other reporting, e.g. AnaCredit. Output is generated by joins, filters, select statements, and by performing simple mathematical operations. Formal joins of counterparty and instrument's tables, selection of subset of variables and filtering out the assets held for trading or held for sale are also formally documented to generate template F 05.01. As shown in Figure 2, calculations are publicly available to market participants.

[3] Bank of Italy Cooperazione PUMA and Central Bank of the Republic of Austria (OeNB) cooperation with Austrian banks. See also Kienecker et al (2018)

[4] Banks' Integrated Reporting Dictionary (BIRD), https://www.ecb.europa.eu/stats/ecb_statistics/co-operation_and_standards/reporting/html/bird_content.en.html

Figure 2. Transformation rules are made publicly available in BIRD

Simplified representation



3. Formats – from technical towards business

Transformation rules that help the interpretation of the reporting requirements may be expressed in different formats: in semantic language, e.g. natural language, or in technical language, e.g. SQL, VTL, Python. In the case of the BIRD project, the Validation and Transformation Language (VTL)[5] was chosen as the formal and technology neutral language to write end-to-end transformations rules. Figure 3 shows an example of transformation rule in VTL for the derivation of the carrying amount depending on the accounting classification[6].

The choice of a technical language typically requires that transformation rules are successfully tested end-to-end with the benefit of ensuring high quality of the transformation rules, e.g. accuracy and consistency, of providing unambiguous interpretation and of encouraging technical implementations. The use of a technical language has therefore several advantages for standardisation.

However, business experts in collaborative initiatives are often required to learn the technical language before they could start writing transformation rules. The limited technical knowledge accompanied with the lack of freely or widely available tooling support may imply that business experts are not in the position to effectively and timely contribute nor to fully benefit from transformation rules. Moreover, end-to-end executable transformation rules are envisaged primarily for testing the quality and for potential implementation by software vendors. In practise however, for performance reasons of each IT environment, testing activities may require custom adaptations of transformation rules, resulting in a multiplication of effort, e.g. transformation rules completely rewritten in other languages, instead of being translated and directly used.

[5] SDMX Technical Working Group, VTL Task Force (2015)

[6] BIRD 5.0, https://www.ecb.europa.eu/stats/ecb_statistics/co-operation_and_standards/reporting/html/bird_content.en.html

Figure 3. Example of VTL transformation rule in the BIRD database

```
define operator D_CRRYNG_AMNT (ACCNTNG_CLSSFCTN component, FV component, GRSS_CRRYNG_AMNT_E_INTRST
component,
                    ACCRD_INTRST component, FV_CHNG_HDG_ACCNTNG component, ACCMLTD_IMPRMNT component,
                    CRRYNG_AMNT component, IS_CRRYNG_AMNT_DRVD component)
  returns component is
    if IS_CRRYNG_AMNT_DRVD = "T"
    then
      if ACCNTNG_CLSSFCTN in {"14", "6", "8", "4", "2", "41"}
      then
        FV
      else
        GRSS_CRRYNG_AMNT_E_INTRST + ACCRD_INTRST - FV_CHNG_HDG_ACCNTNG - ACCMLTD_IMPRMNT
    else
      CRRYNG_AMNT
end operator;
```

Meanings of variables and codes used:
CRRYNG_AMNT: Carrying amount
ACCNTNG_CLSSFCTN: Accounting Classification (
2 = Financial assets held for trading; 6 = Financial assets at amortised cost; 8 = Financial assets at fair value through other comprehensive income; 4 = Financial assets designated at fair value through profit or loss; 14 = Cash balances at central banks and other demand deposits; 41 = Non-trading financial assets mandatorily at fair value through profit or loss)
FV: Fair value
GRSS_CRRYNG_AMNT_E_INTRST: Gross carrying amount excluding interest
ACCRD_INTRST: Accrued Interest
FV_CHNG_HDG_ACCNTNG: Fair value changes due to hedge accounting
IS_CRRYNG_AMNT_DRVD: Is carrying amount derived (parameter to trigger the generation of the Carrying amount)
ACCMLTD_IMPRMNT: Accumulated impairment

Experience had shown that a business component for transformation rules is needed on top of and prior to the technical one. The format of transformation rules needs to be suitable to business experts because they are often those in charge of their creation. Therefore, a design of transformation rules should meet their needs, as "producers" and as "users". On the other hand, the format should also be useful and practical for technical implementation. IT solutions, e.g. in the form of prototypes, are essential to trigger quality control processes on the transformation rules, and to improve the consistency and accuracy of the content, e.g. via multiple decentralised testing processes.

Based on our experience, we conclude that the format of transformation rules should bridge business and technical users. We identified guiding principles for writing standardised and business-friendly transformation rules:

- functional classification and organisation depending on the type of result, e.g. derived attribute, output report
- attention to separation of concerns[7], e.g. logical vs technical, to fit different purposes and users
- the selection of free, open-source and widely known formats to lower the barrier to contribute and to incentivise the development of tooling
- test-driven development approach where tests and testing activities are integral part of the development of transformation rules
- collaborative development and incremental contributions

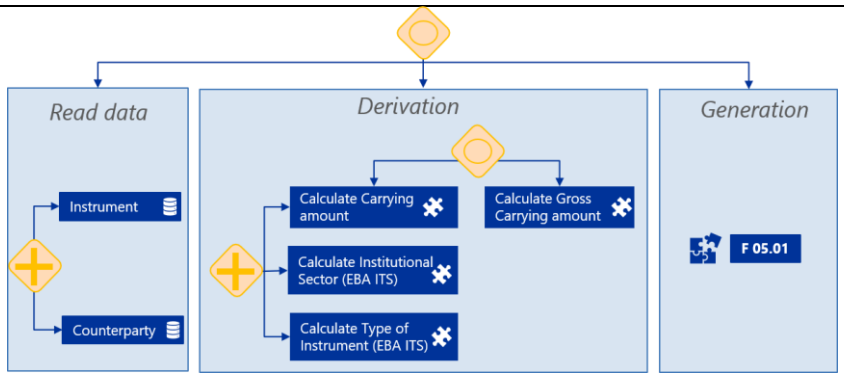
By following the guiding principles, we outline one possible way of writing transformation rules in a collaborative initiative that is favourable to become a standard, e.g. used by business and technical users.

[7] Bos et al (2019)

Business experts provide first an illustration of the flow of transformation rules required from the input to the output using a known standard, e.g. Business Process Model and Notation (BPMN[8]). Such illustration, or diagram, depicts the sequential steps to be followed: starting from reading the input, enriching it with derived concepts, and finally generating the desired output. In our example of F 05.01, business experts sketch the diagram using the agreed convention; in this phase they also identify which concepts are already derived by an existing transformation rule and which need to be derived to feed the output report (Figure 4).

Figure 4. Example of workflow for F 05.01

Simplified representation



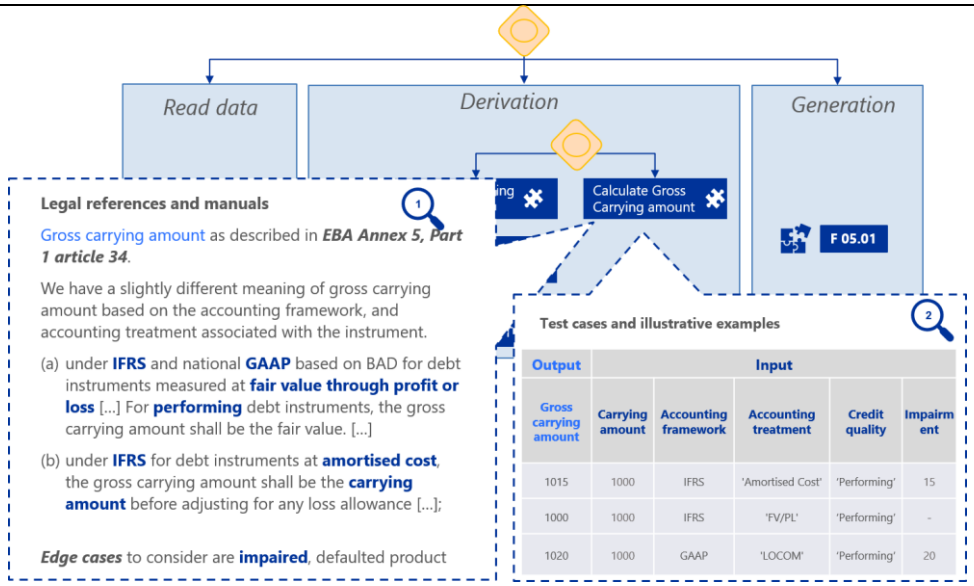
Under this approach, experts and users exploring transformation rules can rely on high-level lineage information and a workflow diagram that shows the order in which operations should be executed. This provides a business-oriented way of exploring a large set of transformation rules at a high level, together with the possibility of retrieving specific details when necessary.

Indeed, the high-level information of the diagram is then enriched with descriptions for each transformation rule. In particular, the business experts fill in, in a pre-defined and business-friendly format, relevant information for business and technical users. For example, a relevant set of information for deriving enriched concepts is: references to legal documents and reporting manuals, explanation of scenarios for illustrative cases, input and output variables, test cases and test data rooted in the same data dictionary, e.g. input and output data models in terms of tables, variables, and allowed values (Figure 5).

[8] https://www.bpmn.org/

Figure 5. Focus on the derivation of gross carrying amount

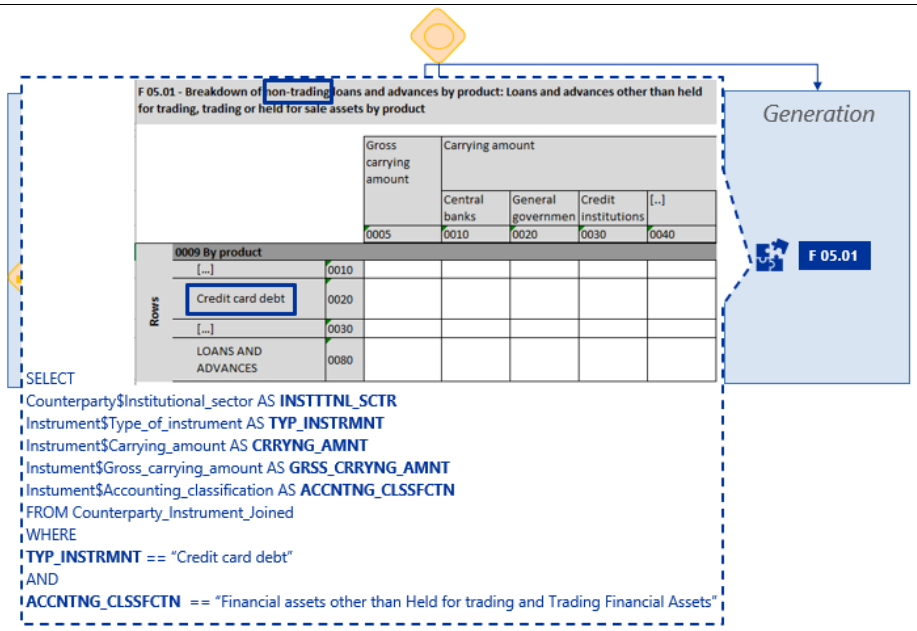
Simplified representation



After all derived concepts are defined, business experts finalise the documentation of transformation rules for each desired output, e.g. they define which aggregation, grouping or filtering is needed to match the requirements expected for template F 05.01. The definition of steps for the generation of output requires simple mathematical operations which are therefore well documentable by non-technical experts also using a programming language, e.g. SQL, VTL, Python. The intention is to enable business experts to contribute time-effectively and with business value added, and not to write a fully testable and often not business-friendly piece of code. This can be achieved by accepting a simplification step, i.e. by assuming that correct join statements across all required input tables are applied and create a single flat table, thereby reducing the complexity of steps that business experts need to document (Figure 6).

Figure 6. Focus on the generation of F 05.01

Simplified representation



In this example, business experts do not create end-to-end executable transformation rules. They create all the business artefacts required for software vendors or reporting agents' IT developers for implementing an IT solution that uses the technology they find more suitable. Experts can carry out their contribution to the definition of transformation rules by analysing regulations and by using accessible tools and formats to write down their findings. The output is then published and available not only to other business experts for consultation and for potential validation and improvement, but also to technical experts, e.g. in the field of software development, for further enrichment and application, e.g. into an IT solution.

4. Conclusions

The field of regulatory reporting shows a high degree of heterogeneity of formats and semantical description, hence standardisation on how requirements are documented is constantly under authorities' consideration[9]. Possible gains for cost reduction and data quality can be achieved in the future with a European integration of requirements[10] and a standardisation of formats. By means of standardised transformation rules, a considerable burden on financial institutions and software vendors is reduced and resources can be freed for more competitive tasks. Additionally, a standardised format with a focus on business language is important for a wide contribution, adoption and as a basis for innovation.

Our conclusions point to the need of designing a standardised and business-friendly format for business transformation rules prior to any documentation in technical language. Such solution overcomes barriers and long learning curves that may hinder effective contribution, and offers the market participants an environment that is favourable for experimenting and testing prototypes.

Future initiatives on standardisation in regulatory reporting can vastly benefit from a collaborative development of transformation rules. In this context, we agree that the two pillars of the ESCB long-term strategy for banks' data reporting, BIRD and IReF, are an opportunity of a playground at an early stage to continue developing synergies in aligning data models[11]. Such synergies could be exploited in the future for what concerns standardised transformation rules, and therefore we identified principles to design a solution that bridges business and technical needs and creates a favourable requisite for wide adoption.

While a fully integrated reporting is a long-term objective, collaborative initiatives have the possibility to proceed at faster pace. They can absorb good national and international practices, set aside identified deficiencies of existing approaches, and step up their engagement on leading conceptual work and consolidation of practical experiences in the design of transformation rules.

References

Banks' Integrated Reporting Dictionary, https://www.ecb.europa.eu/stats/ecb_statistics/co-operation_and_standards/reporting/html/bird_dedicated.en.html

Bos, A. and van der Helm, R. (2019): "Heading for harmonization of data collection"

[9] Crisanto et al (2020)

[10] European Central Bank (2021)

[11] European Central Bank (2021), Colangelo et al (2021)

Business Process Model and Notation, https://www.bpmn.org/

Crisanto, J.C., Kienecker, K., Prenio, J. and E. Tan (2020): "From data reporting to data-sharing: how far can suptech and other innovations challenge the status quo of regulatory reporting?", FSI Insights on policy implementation, no 29, December.

Colangelo, A., Gross, F. and Schuster, F. (2021): "Effective measurement of the economy in the emerging digital age"

European Banking Federation (2019): "Boosting Europe: Building Trust and Supporting Growth in Europe. EBF recommendations for the EU 2019-2024 legislative cycle and beyond", May

European Central Bank (2021): The Eurosystem Integrated Reporting Framework: an overview, December

Kienecker, K, G Sedlacek and J Turner (2018): "Managing the processing chain from banks' source data to statistical and regulatory reports in Austria", Statistiken – Daten und Analysen, August.

PUMA, https://www.cooperazionepuma.org/homepage/index.html

SDMX Technical Working Group, VTL Task Force (2015): "Validation & Transformation Language, Part 1 – General description. Version 1.0"

DEFINING BUSINESS TRANSFORMATION RULES IN A STANDARDISED FORMAT*

A PRACTICAL CASE

IFC-BANK OF ITALY WORKSHOP ON "DATA SCIENCE IN CENTRAL BANKING"

17/02/2022

ARRU D., ODDONE G.

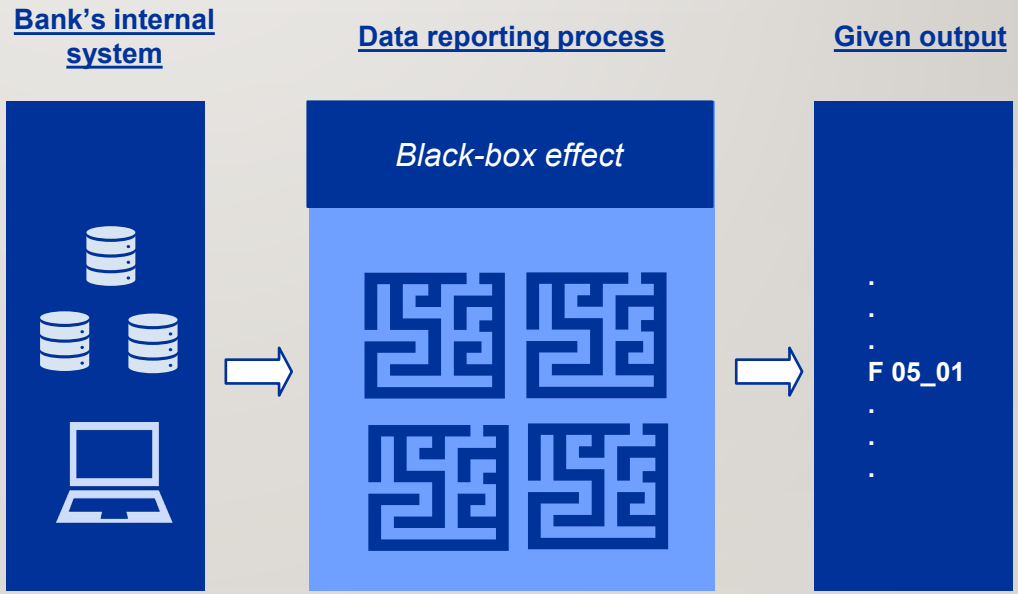
* THE VIEWS EXPRESSED ARE SOLELY THOSE OF THE AUTHORS AND DO NOT NECESSARILY REFLECT THE OPINION OF THE EUROPEAN CENTRAL BANK

In the regulatory reporting fields scattered lack of standardisation has multiple consequences:

- → Interpretation of regulations is a highly intense work for banks and software vendors
- → Reporting processes are all different, customised to banks, and hard to change

Question: What can be done about it?

Example: to produce the same Finrep F 05_01 template, each bank utilises different input models and follows different data reporting processes

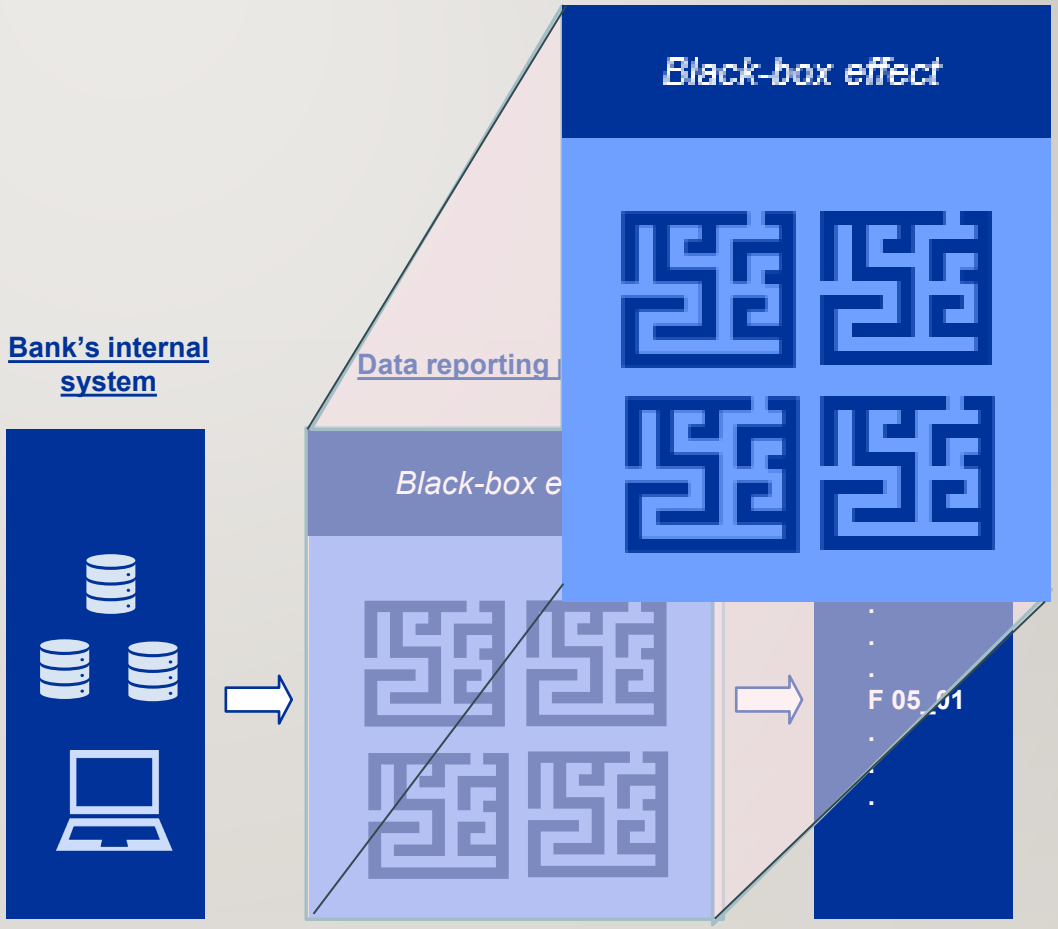


Transformation rules in data reporting processes play an important role in standardisation effort, since they describe operations to derive information and create reports.

The choice of the format of these descriptions is a key element to move away from the black-box effect.

Less technical and more business-friendly:

- ✓ More business contribution
- ✓ More usable and adoptable
- ✓ Less developing and testing



Let's see an everyday example…

Instructions to build wardrobes are highly standardised and we all benefit!

And when they are not following a standardised (and user-friendly) format, we as users experience difficulties to interpret them, e.g.:

Do we have all the pieces?

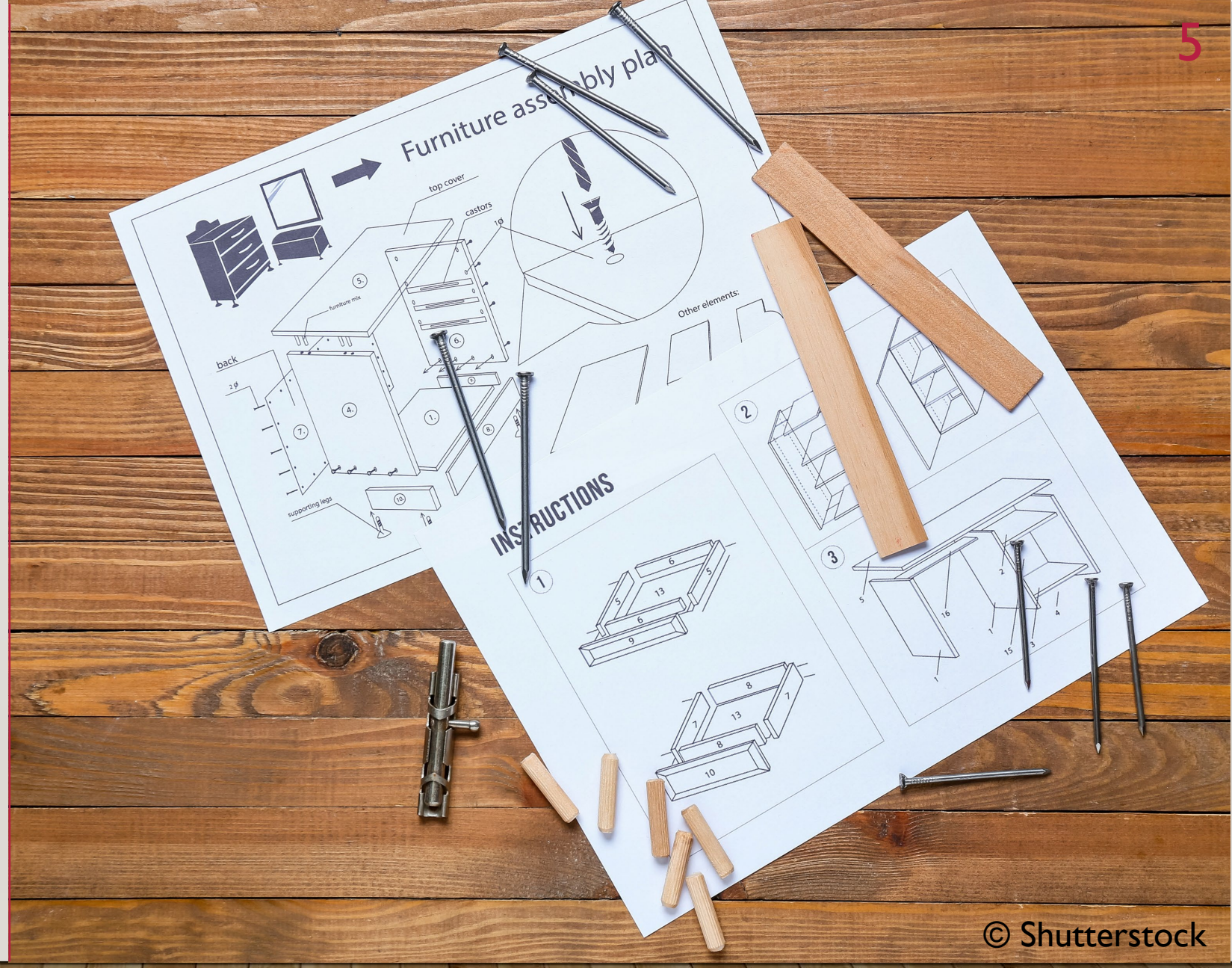
When do we assemble the drawer?

Before or after the mirror?

A standardised and user-friendly format for instructions leads to clarity and provide the basis for further improvements.

Likewise a standardised and business-friendly format for transformation rules offers many benefits:

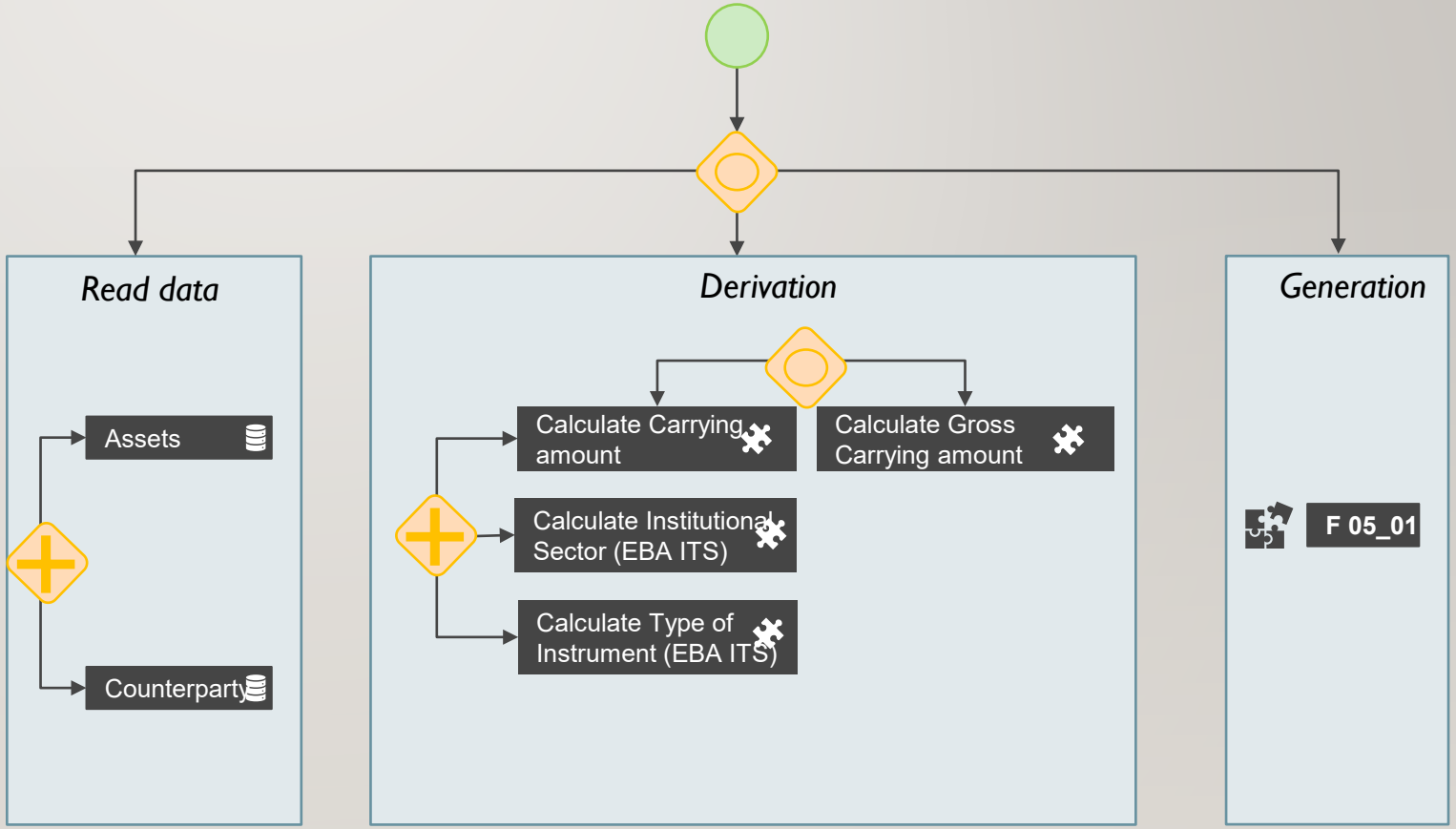
- ✓ adopted by market participants
- ✓ understood by experts
- ✓ further improved by technical experts
- ✓ IT solutions based on it



Principles for a standardised and business-friendly transformation rules to bring standardisation and incentivise innovation:

- ✓ Publicly documented in a collaboration effort between authorities and banks
- ✓ Respect the principle of separating concerns to better fit needs and focus contribution
- ✓ Founded on open-source and known formats for wider tools support
- ✓ Suitable basis for technical implementation

Example: to produce the same Finrep F 05_01 template, each bank has access to publicly available transformation rules organised and documented in business-friendly and standardised formats.



The choice of format for transformation rules should fit the needs of different interested parties and is an important element to feed in future initiatives.

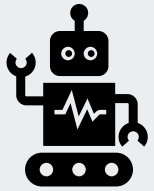
It facilitates contributions by:

- ✓ Internal business contributors
- ✓ External contributors

It provides useful descriptions and instructions of reporting requirements to:

- ✓ Business users
- ✓ Technical profiles

Example: standardised and business-friendly transformation rules for multiple benefits and different needs

	✓ Legal references and manuals
	✓ Diagram of the data transformation process
	✓ Attribute lineage information
	✓ Illustrative examples
	✓ Test cases
	✓ Business artefacts required for implementation
	✓ Technical format

Some conclusions:

- Standardisation in the field of reporting and data processing is beneficial (and needed)
- Transformation rules can be documented in many formats and languages, e.g. natural, business, technical language
- The definition of a format can have relevant impact on future initiatives around reporting burden and integrated reporting requirements
- A standardised and business-friendly format for transformation rules improves the successfulness of the standardisation efforts. E.g. clarity for business experts to contribute and use, basis for technical experts to develop and improve