CGAP

WORKING PAPER

Data Collection by Supervisors of Digital Financial Services

Denise Dias and Stefan Staschen

ABSTRACT

Reliable, timely, complete, and readily accessible data are crucial for efficient and effective risk identification and assessment in financial sector supervision and enforcement. They are particularly important for financial supervisors in emerging markets and developing economies (EMDE) who face fast innovation and a regulatory perimeter that is getting bigger because of growing digital financial services (DFS) for financial inclusion in EMDE. Which data should the DFS supervisor collect? How frequently? In what format? Through which means? How should the DFS supervisor improve data? What aspects should be considered? Based on primary and secondary research, this working paper highlights practices in data collection for DFS supervision and the issues and emerging trends in regulatory data collection mechanisms that are relevant to DFS supervisors. It provides material for DFS supervisors to exploit opportunities created by developments such as RegTech and SupTech to rethink their approach to data collection, with the goal of strengthening supervision while fostering financial inclusion.

ACKNOWLEDGMENTS

Consultative Group to Assist the Poor/World Bank Group

LIST OF ACRONYMS

AML/CFT	Anti-money laundering and combating the financing of terrorism
ATM	Automated teller machine
BNR	National Bank of Rwanda
BoT	Bank of Tanzania
CBN	Central Bank of Nigeria
CNBV	Mexico's National Banking and Securities Commission
Consar	Mexico's National Commission of the Pension System
CPMI	Committee on Payments and Market Infrastructures
CRD	Document Submission Control system by the Central Bank of Brazil
CSSF	Luxembourg's Commission de Surveillance du Secteur Financier
DFS	Digital financial services
EBA	European Banking Authority
EMDE	Emerging markets and developing economies
EU	European Union
FCA	U.K.'s Financial Conduct Authority
G2P	Government-to-person
GPS	Global positioning system
KYC	Know your customer
MNO	Mobile network operators
OeNB	Austria's central bank (Österreichische Nationalbank)
OFR	U.S. Office of Financial Research
P2P	Peer-to-peer (or person-to-person)
POS	Point-of-sale
RegTech	Regulatory technology
SBP	State Bank of Pakistan
STR	Suspicious transaction report
SupTech	Supervisory technology
TCRA	Tanzania Communications Regulatory Authority

TABLE OF CONTENTS

LIST OF TABLES, FIGURES, AND BOXES

INTRODUCTION

"[O]versight of financial markets and policy decisions are only as good as the data supporting them" (OFR 2016).

Financial supervisors' approach to data can have profound implications on the effectiveness of their work—and collecting good data is not always easy. Data collection can be costly for supervisors and reporting institutions, particularly when the supervisor's data needs change constantly to reflect market, regulatory, and supervisory shifts. In 2003, the Basel Committee for Banking Supervision noted an "unprecedented speed of change related to technological and customer service innovation, the ubiquitous and global nature of open electronic networks, the integration of e-banking applications to legacy computer systems and the increasing dependence of banks on third parties." This description could apply to today's world, but now nonbanks play a much more relevant role in a makeover of financial services. Relentless innovation in digital financial services (DFS) (see Box 1) is seen everywhere across advanced economies and emerging markets and developing economies (EMDE),[1] both as a cause for and an outcome of reforms such as new

regulations on agents, e-money, financial technology (FinTech), consumer and data protection, and cybersecurity. These developments have an impact on what data financial supervisors need and on the mechanisms used to collect the data.

EMDE supervisors may find it more challenging to deal with DFS innovation than would supervisors in advanced economies. This may be because of (*i*) thinner regulatory frameworks; (*ii*) the higher significance of nonbanks involved in DFS and business models that do not fit existing regulations; (*iii*) the emergence of products and services that blur the lines between banking, securities, payments, insurance, and nonfinancial sectors; (*iv*) the fact that DFS may be the first and only contact of millions of (often lower-income) customers with the formal financial sector; and (*v*) limited supervisory capacity, skills, and resources, including adequate IT systems.

In response to these challenges, EMDE supervisors may make extra efforts to

---

> ### BOX 1. Definition of *digital financial services* for this research
>
> DFS is often broadly defined to include any type of financial service or product delivered through any digital device and digital transactional platform, and provided by any type of bank or nonbank institution. This encompassing definition would be adequate for many aspects covered in this paper. However, the intent of this paper is to offer ideas for DFS supervisors in EMDE, who want to foster financial inclusion through DFS. In a financial inclusion context, DFS usually deploy widely dispersed agent networks and are often provided or facilitated by nonbanks such as e-money issuers. This research focuses on the data collected by supervisors of nonbank e-money issuers and their e-money operations, the agents used by nonbanks and banks, and the statistics collected to support payments oversight by central banks.

---

1 As per country classification by the World Economic Outlook 2017 of the International Monetary Fund. See http://www.imf.org/external/pubs/ft/weo/2017/01/weodata/weoselagr.aspx.

> A data collection mechanism is the combination of the systems and procedures used to compile, transform, validate, and report data (reporting institution's side), and the systems and procedures used to collect, validate, store, manage, and access the data (supervisor's side).

increase the amount of data collected on DFS. Our research confirms that several EMDE supervisors are collecting increasingly larger volumes of more granular data, including DFS data. The fast development of regulatory technology (RegTech) and supervisory technology (SupTech) supports and encourages these efforts and paves the way for an emerging and broader shift in regulatory and supervisory approaches based on better data and experimentation.[2] Some of the challenges EMDE supervisors currently face may become a strength: the lack of legacy systems and extensive regulatory frameworks may make it easier for EMDE supervisors to adopt and encourage adoption of RegTech and SupTech solutions to leapfrog into an era of data-driven regulation and supervision.

This paper provides insights into two closely related dimensions of data collection for DFS supervision: the specific data required and the data collection mechanisms used by supervisors to gather such data. Our research maps reporting requirements in a few key areas of DFS in a sample of countries. The paper also addresses the more complex and relevant discussion of underlying data collection mechanisms and highlights important issues that can affect the efforts by supervisors to improve DFS data. This discussion and the examples used are not limited to DFS-specific data

collection mechanisms, because practices and systems used to collect any type of regulatory data offer relevant lessons for DFS data collection. Also, DFS is not an isolated item of regulatory data collection mechanisms that requires special treatment. Efforts to improve DFS data are an integral part of the overarching approach to data collection at financial supervisory authorities.

This paper reflects a combination of primary and secondary research that were conducted from mid-2016 through mid-2017. The primary research comprised in-person and phone interviews, as well as questionnaires filled out by some supervisors (see Annex 1 and Annex 2). The interviews were the main source of information on the practices of individual countries. The secondary research comprised an extensive desk review that included relevant international standards and guidance, third-party research and news articles (see Annex 3), and regulatory reporting requirements for nearly 40 jurisdictions. The requirements researched included data collected on e-money issuers, use of agents, and electronic retail payment statistics.

Data collected on e-money issuers excluded financial products that are provided by third parties that use e-money issuers' platforms as a delivery channel, such as digital credit, mobile insurance, pensions, and savings.[3] For this research,

---

2 For discussions of the role of data and experimentation in emerging regulatory and supervisory approaches, see Zetzsche et al. (2017) and Arner, Barberi, and Buckley (2010). See Viñals et al. (2010) and BCBS (2015) for overviews of how financial supervision is changing globally toward a more intrusive supervision.

3 Many nonbanks in EMDE offer mobile money, which is a stored-value product, or e-money delivered via mobile phones. These nonbanks are called *mobile money providers* (i.e., a type of e-money issuer). Mobile money providers can establish partnerships with other financial institutions to distribute savings, insurance, pensions, and investments products on their behalf, on their mobile money platforms.

the term *e-money issuers* includes the following:

- Dedicated nonbank e-money issuers.

- Dedicated nonbank e-money issuers that are not directly regulated but that operate under the auspices of a licensed bank in bank-led models (e.g., Bangladesh, Myanmar).

- Nonbank e-money issuers that are not dedicated to financial services (e.g., mobile network operators [MNOs] that issue e-money).

- Limited-purpose banks that are dedicated to issuing e-money and providing payment services (e.g., payment banks in India and niche payment banks in Mexico).

For this research, *agents* includes e-money agents and agents used by banks and other regulated financial institutions in what is often referred to as *agent banking*. The term, as used in this paper, does not include insurance agents. In the case of Mexico, a few practices by the pensions regulator Consar, regarding agents of pension administrators, were cited when they were relevant to DFS data collection.

In this paper, *electronic retail payment statistics* refers to data traditionally collected by payments system oversight departments or units at central banks. These data include volumes and value of e-payment transactions and payment instruments (e.g., e-money accounts, debit and credit cards).

The sample of countries used in this research comprises leading EMDE markets where DFS for financial inclusion have achieved scale or are growing rapidly. The sample also included several developed economies for comparison purposes and EMDE that have either or both agent regulations and e-money regulations. The European Union (EU) was included in the research of regulatory reporting requirements and in the discussion of data collection mechanisms, because of its wealth of relevant material and the potential application of certain practices to DFS supervision in EMDE. The authorities interviewed or that have filled out questionnaires held primary responsibility for supervision of the institutions that offer DFS (central banks and/or supervisory authorities). In Mexico, Consar was also included given the relevance of its data collection practices to the discussions in this paper.[4] A few other institutions, such as DFS providers and providers of IT and consulting solutions, were also interviewed.

This paper intends to help deepen the knowledge of the practices, challenges, and trends in regulatory data collection in general and for DFS specifically by (*i*) discussing the basics in regulatory data collection and illustrating practices through country examples, (*ii*) identifying the range of DFS data required in researched countries, and (*iii*) identifying emerging trends in regulatory data collection that could affect DFS data collection and supervision. This paper highlights the pros and cons of some practices and opportunities created by technology. It is not a compilation of best practices, nor is it an assessment of any country approach. The paper does not make specific recommendations as to which types of data collection mechanisms should be adopted by DFS supervisors in EMDE and the content of the data that need to be collected by DFS supervisors.

---

4   Unless Consar is expressly mentioned, all other references to Mexico in this paper relate to the regulations and practices by the National Banking and Securities Commission (CNBV).

The paper builds on and complements previous and forthcoming CGAP work on developments in DFS markets and DFS policy, regulation, and supervision, including ongoing work on the potential impact of RegTech on financial inclusion. The primary audience for this paper is EMDE supervisors that are responsible for DFS, particularly those that oversee nonbank e-money issuers and the use of agents by regulated institutions. Because the paper includes aspects generally applicable to any regulatory data collection, it may be relevant to other EMDE financial supervisors.

The research described in this paper has several limitations: the mapping of DFS reporting requirements faced some challenges (highlighted in Section 3) and the research does not cover some important issues such as (*i*) how DFS data are used by supervisors and the challenges they face (a forthcoming CGAP publication will address supervisory procedures for e-money issuers),

(*ii*) the costs of different data collection mechanisms (some enquiry was done but there no conclusive findings), (*iii*) the availability of IT solutions and vendors across countries, (*iv*) how institutional and supervisory culture shape data collection mechanisms, and (*v*) data collected on digital credit, mobile insurance, mobile savings, and mobile pensions when distributed, but not directly provided, by e-money issuers (see Box 4). Finally, further research is needed to shed light on the critical issue of supervisory capacity and data analytics skills at supervisory authorities in EMDE, since improving DFS (and other) data will have limited impact if capacity and skills are lacking.

A Glossary of terms used in this paper is found in Annex 4 and a few key terms are explained in Section 2.1. In addition, several examples of documents used by supervisory authorities to implement DFS (or other) reporting requirements are found in CGAP (2017).

1. WHAT ARE GOOD DATA?

Good data are at the heart of any effective financial supervision, including DFS supervision. However, even supervisors that have good data may face difficulties due to, for instance, lack of analytical capacity or human resources, so quality data are the minimum one could ask for. What are good data?

Literature on general data management points to three attributes of completeness, timeliness, and accuracy (see Figure 1). These elements are easier to understand through illustrative examples of what it means when they are lacking.

Case 1. Through the newspapers, the DFS supervisor learns about a major fraud case where employees of a nonbank e-money issuer tampered with the account management system to steal customers' funds. Specifically, the fraudsters manipulated *suspense accounts*. Despite receiving daily returns from all e-money issuers, the supervisor had no information on suspense accounts. The data are **incomplete**.

Case 2. The e-money industry in Emoneyland is growing at a pace of 300 percent per year, but the e-money issuers report only every April to the DFS supervisor about their operations in the previous year. The data are **not timely**.

Case 3. A monthly regulatory report from an e-money issuer informs that the total e-money issued is US$10 million. When inspecting the e-money issuer, the supervisor finds that the total e-money issued is actually US$100 million. The reported data were **inaccurate**.

A range of factors—including ill-designed report templates or unclear purposes for the data collected—could lead to inaccurate, incomplete, or untimely data. Our research indicates that certain problems could be avoided by using better data collection mechanisms.[5]

FIGURE 1. Three key attributes of good data

Complete	Timely	Accurate
Nothing is missing, and data fulfill the supervisor's data needs.	Data come at the right time and frequency, and are not excessively outdated.	Data are reliable, credible, and correct.

---

5   This finding is corroborated by previous publications (e.g., OFR 2015, 72 and 74; IIF 2016, 8).

2. WHAT DFS DATA ARE NEEDED?

2.1 Terminology

Some terms related to regulatory reporting are used repeatedly throughout this paper, and their meanings may differ depending on the source. A full list of terms and the definitions used for this publication can be found in the Glossary, but a few key terms are also presented in tables 1 and 2.

Data can be quantitative or qualitative, and financial or operational (see examples in Table 2). Periodic reporting of DFS data in the countries that were researched focuses on operational data, while financial data, such as financial statements, are collected less frequently.

2.2 Mapping data needs to achieve policy and supervisory objectives

Mapping data needs should be the first step in data collection, but not all supervisors who were interviewed had formally done so before creating DFS reporting requirements. Also, not all supervisors had collaborated across departments or authorities to identify common needs. The result, in some cases, was incomplete and/or duplicate reporting requirements.

The type of information supervisors need depends on several factors, such as how policy objectives (e.g., keeping financial systems stable, fostering financial inclusion, protecting consumers, curbing financial crime, and increasing efficiency in payment systems) are translated into specific supervisory objectives. While many supervisors may have similar policy objectives, they differ on their supervisory objectives and actions (BCBS 2015), therefore, the information they need and collect will also differ (see Section 3). For instance, in the researched advanced economies where nonbank e-money issuers are not

a priority for supervision because of their lower relevance compared to banks, less data are collected on them than in sample EMDE that had a substantial nonbank e-money issuing industry that serves millions of customers. Also, it seems that at least part of DFS data in EMDE (e.g., geographical location of agents) are collected to monitor financial inclusion indicators, which may be considered less relevant to advanced economies.

By identifying the information needed for each supervisory objective, authorities can precisely define elements of their reporting requirements: the **scope** (the specific content to be collected), **periodicity** (the frequency with which data are collected), and the **format** (level of aggregation and the definitions to be used). Table 3 illustrates data needs mapping. It shows how different policy objectives can have overlapping information needs.

Mapping data needs helps authorities identify when the same data points will be needed by different departments or authorities. Mapping data helps to determine who is responsible for collecting (and sharing) the data. This could reduce duplication in reporting requirements, including when authorities outside the financial sector are involved. For instance, nonbank mobile money providers in Tanzania report mobile money transactions to the Tanzania Communications Regulatory Authority (TCRA), because the parent companies are MNOs. They also report transaction data (in a different format) monthly to the Bank of Tanzania (BoT). TCRA and BoT are working together to resolve this duplication.

The mapping process also helps to determine which institution should report which data. For instance, if information on digital credit and insurance distributed through e-money issuers is needed (e.g., for prudential supervision

TABLE 1. Definition of key terms, as used in this paper

Term	Definition
Regulatory report or return	A formal report, prepared according to reporting requirements previously defined by the authority, that the supervisory authority receives from institutions. The report may have aggregated data or granular data (or both).
Report template	A model regulatory report created by supervisory authorities. Reporting institutions need to fill in the template using manual or automated processes (or both). A template has several data fields that contain aggregated or granular data, in raw or standardized format.
Data field	Each data field in the report template needs to be filled in. Data fields can include one or more data points in raw or standardized format.
Data point	Specific data that need to be provided. Sometimes several data points are combined through formulas to fill in a data field. Data points in DFS reporting commonly represent aggregated data (e.g., total number of agents) but also can be granular data.
Aggregated versus granular data	DFS reporting requirements are usually high-level aggregated data, such as *total outstanding e-money issued* and *total value of transactions*, which are calculated through formulas defined by the supervisory authority or by the institution (e.g., total outstanding e-money issued equals the sum of all e-money client account balances at the end of the day prior to the reporting date, after all fees are deducted). Aggregated data can also be reported as an indicator, such as percent *of active e-money accounts*. Granular data are presented in a higher level of detail. There are different levels of granularity. At the most granular level, data are similar to those that are generated on an ongoing basis by financial institutions in the course of their businesses (e.g., client tables, transaction tables, loan tables, access logs, etc.). Requiring a list of all agents and all transactions per agent, or a list of all e-money transactions, is considered granular data, even though this type of reporting requires some data aggregation. Both aggregated and granular data can be reported in raw or standardized formats.
Standardized versus raw data	Raw (or business) data are generated by the information systems of a reporting institution. Since institutions use systems that comply with the specifications of the IT vendor and/or the institution, raw data are usually not comparable across institutions, raw data are generated at the most granular level (e.g., client tables, transaction tables), and they also can be (and usually are) aggregated at some level for internal management purposes (e.g., risk management). Standardized data are raw data that have been put into a common format that is defined by the supervisory authority, including by using common definitions and formulas. Most data collected by financial supervisory authorities are standardized, but there are examples of collection of raw data for DFS supervision (both aggregated and granular).
Data category and subcategories	Data category refers to high-level fields in reporting templates that may be broken down into other data fields called subcategories. An example is *total volume of transactions*. This category can have several transaction types as subcategories, such as *transfers*, *deposits*, *withdrawals*, etc.

or monitoring of financial inclusion indicators), then such information would need to be reported by the lenders and insurers, not by the nonbank e-money issuer. The lender and the insurer are the legal providers of the services and hold the needed information, while the e-money issuer is only a distributer of the services.[6] In at least a couple of researched countries, nonbank

TABLE 2. Examples of types of data collected by DFS supervisors

	Financial	Operational
Quantitative	Financial statements (balance sheet, cash flow, income statement) Financial ratios (capital adequacy ratio, liquidity ratio, and others)	• Volume and value of transactions • Number of transaction points (e.g., agents, ATM, point of sale) • Number of accounts and total balances • Losses from frauds, consumer compensations
Qualitative	Explanations to financial statements	• Policies and procedures • Description of frauds and actions taken, actions taken on consumer complaints, IT systems, risk management practices, accounts of service disruptions

TABLE 3. Example of how data needs can be mapped by financial authorities

Examples of **policy objective**[a]	Increase efficiency of the national payments system	Protect financial consumers
Example of specific **supervisory objective**	Monitor the evolution of the use of electronic means of payment	Monitor the evolution of consumer complaints related to nonbank e-money issuers
Example of data need (**scope—high level**)	Indicator 1: penetration of e-money accounts Indicator 2: use of e-money accounts	Indicator A: number of complaints about e-money issuers relative to the number of e-money accounts issued
Example of data need (**scope—data points**)	Indicator 1: (*i*) total number of e-money accounts; (*ii*) total adult population Indicator 2: total volume of merchant purchase transactions in e-money accounts	Indicator A: (*i*) total number of e-money accounts, by e-money issuer; (*ii*) total number of complaints, by e-money issuer
Example of data need (**periodicity**)	Indicator 1: semi-annually Indicator 2: semi-annually	Indicator A: quarterly
Example of data need (**format**)	Aggregated data (e.g., reporting will be of indicators 1 and 2 calculated and reported by the institution)	Aggregated data (e.g., reporting will be of indicator A calculated and reported by the institution)

a. Note that for each policy objective there will be numerous supervisory objectives. For each supervisory objective, there will be numerous indicators that can be broken down into numerous data points. This table gives examples of each for illustration purposes only.

---

6   Most country regulations on nonbank e-money issuers prohibit the issuers from offering credit and insurance. When they distribute such services, they do so on behalf of the legal providers of the services, which are usually the bank (or other type of lender) and an insurance company. The nonbank e-money issuer is merely a distributor, it does not assume credit and insurance risks and does not hold data on the credit and insurance contracts.

e-money issuers are required to report digital credit and insurance data (e.g., number of insurance policies, number of loans). This situation raises secrecy concerns (because a third party—the e-money issuer—may access data on credit and insurance clients), and it is likely to mean duplicate reporting done by the lender and insurer to their prudential supervisors. Another situation found in a few researched countries was the supervisor requesting the bank holding the funds backing the e-money issued (the float/trust account) to report on e-money transactions, even though the bank is not the e-money issuer.

Carefully mapping data needs may reduce compliance costs for reporting institutions. Some supervisory authorities may in fact be required by law or otherwise to keep compliance costs down. This includes the United Kingdom's Financial Conduct Authority (FCA), which must estimate compliance costs of reporting requirements, and show that the benefit justifies the costs. Supervisors in the EU are required to reduce the reporting burden.[7] In 2013 the Central Bank of Brazil instituted a program, Otimiza BC, to increase the efficiency of its operational and administrative processes, including rationalizing reporting requirements.

Keeping compliance costs down is all the more important for DFS in a financial inclusion context, and for small institutions. For these, the relative cost of dedicating staff time to regulatory reporting can be higher than for large institutions and for institutions that have wider profit margins. Accordingly, some researched countries have lower or no reporting requirements for small nonbank e-money issuers. However, regulatory reporting costs (including the costs of changing

reporting requirements) may depend more on the data collection mechanism than on the amount of data, as noted by some interviewees. Reporting a small amount of data that need to be aggregated and formatted into separate report templates can be more costly than reporting a larger quantity of granular data through an automated process that does not use report templates. For more on data collection mechanisms, see sections 6 and 7.

2.3 A standard process for creating reporting requirements

The process to create DFS reporting requirements varies across the authorities that were interviewed—from an informal and unstructured process limited to the individual department that requests the data, to formal, standardized procedures that involve coordination among multiple departments and authorities and consultation with the industry. The authorities from Austria, Brazil, Colombia, Hong Kong, Luxembourg, Malaysia, Mexico, Peru, the Philippines, Singapore, and the United Kingdom, for instance, have codified standard steps and principles for creating reporting requirements that apply to changing/adding/removing items from existing report templates and creating new ones.[8]

A formal standard process helps to ensure data standardization, avoid duplication, and increase quality, transparency, and accountability. Extensive consultations, including testing periods for changes in or new reporting requirements, allow the parties to address challenges and agree on an exact definition of all data points and data fields in report templates, which lessens the potential for misinterpretation.

---

7 See *Commission Implementing Regulation (EU) No. 680/2014.* Also, the *Payment System Directive 2015* directs member states to limit information requirements for low-value payment services providers. See also EBA (2017).

8 In the EU, all supervisors under the jurisdiction of the European Banking Authority have established well-documented processes for implementing updates to the reporting framework (EBA 2017).

BOX 2. Example of a process for implementing reporting requirements

- Based on broad policy objectives, identify specific supervisory objectives that require data.

- For each supervisory objective, identify the information needed and its uses and users.

- Identify the underlying data points that need to be collected.

- Identify whether the data points are already collected across the supervisory authority.

- Draft the report template with standard definitions for each data point and data field.

- Coordinate with other departments and/or authorities to
  a. Avoid duplicate requirements
  b. Agree on key definitions
  c. Assign responsibility for collecting and sharing the data
  d. Secure IT resources to implement the new requirement

- Submit draft report template for external consultation and/or provide a testing period for the new requirement to
  a. Identify significant obstacles to implementation
  b. Consider compliance costs, particularly if implementation requires changes in operational procedures and IT systems
  c. Agree on a realistic implementation timeline

- Monitor quality of the data and fix problems with the template

- Enforce full implementation of the requirement

Because of this, supervisors may spend less post-reporting time with data cleaning/correction. Box 2 gives an example of a formal process for creating/changing reporting requirements.

The process is affected by whether regulatory data collection is centralized or decentralized at supervisory authorities. Several researched authorities (e.g., Austria, Bangladesh, Brazil, Colombia, Ghana, Hong Kong, India, Luxembourg, Mexico, Nigeria, Peru, the Philippines, Rwanda, Singapore, United Kingdom) have a unit that functions as a central data source for all other departments. This unit is responsible for prioritizing and implementing all periodic electronic data collection, including DFS data. A central unit acts as the interface between different departments when creating/changing reporting requirements and when implementing improvements in the data collection mechanism. Keeping a single catalog of all regulatory reports imposed on all types of institutions regulated by the supervisory authority helps to avoid duplicate requirements, which could help in estimating the total reporting burden.

Testing periods are particularly important when changing report templates, creating new requirements, or changing

the data dictionary (i.e., the definitions of all data fields and data points) and the taxonomy (i.e., rules such as the interrelationship between different data fields in a template) or validation rules (see Section 6.2). Testing periods help both sides set realistic implementation timelines, because seemingly minor changes can have substantial practical consequences for reporting institutions. Depending on the case, supervisors may consider setting a cut-off date and requiring new data only for new customers/accounts.[9] The Bank of Ghana, for instance, has recently added *client gender* as a data field in the monthly report template for nonbank e-money issuers. Because the institutions did not have these data previously, they had to gather gender data from all existing customers, change account-opening procedures/materials to collect gender data of new customers, and change their IT systems to add *gender* in client files. Requiring GPS location of agents or transactions would need to involve a similar process.

In some of the countries in this study, the process for creating reporting requirements is part of a comprehensive data governance policy (see the example in CGAP (2017). Data governance is understood to be the management of the availability, usability, integrity, and security of data, which can help supervisors improve the process of creating and implementing reporting requirements. A data governance policy defines the organizational structure and systems for collecting, managing, and disseminating regulatory data; the process for creating or changing reporting requirements; the responsibilities of staff members; the rules for accessing, altering, and deleting data; audit trails; the rules for sharing internally and externally; and overarching principles.

---

9   The Institute of International Finance has identified tight regulatory deadlines for IT updates as a reason why financial institutions end up tinkering around the edges of existing IT infrastructure, instead of undergoing a necessary and more fundamental overhaul of their systems or implementing technological innovations that could improve regulatory reporting. See IIF (2016).

3. DFS DATA IN THE COUNTRIES STUDIED

This section describes the data collected on e-money issuers, bank and nonbank agents, and electronic retail payments statistics (see definitions in the [Introduction](#)). It is important to note the limitations of the findings in this section. The research comprises desk-based analysis of reporting requirements described in regulations and publicly available report templates and statistics. Actual report templates used by DFS supervisors often expand on the regulatory text. (Examples of report templates are in CGAP [2017].) Since only a minority of countries publish report templates, this section may not give a completely accurate view of the data collected in the sample countries. Because of this limitation, this section does not include a country-comparison table of reporting requirements, with details on data categories (e.g., *volume/value of e-money transactions*) and breakdown into subcategories (e.g., *volume/value of e-money transactions by types of transaction*). Instead, it focuses on highlighting common practices and other relevant findings.

The terms used for similar data vary widely across researched countries. The disparity is especially accentuated in the subcategories, particularly the types of e-money transactions. Hence, the terms used in this section are not necessarily those used in each country.

3.1 General observations

➤ Authorities in advanced economies and some EMDE such as Brazil, Colombia, Kenya, Mexico, and Peru publicize in more detail, on their websites, the DFS [reporting requirements](#) than other countries in the research.

➤ There is no wide variation in the broad data categories used for DFS reporting (common categories include, for instance, *volume/value of e-money transactions* and *number of agents*, but there is a very wide variation on the subcategories, especially of transaction types (e.g., *withdrawals*, *transfers*, *payments*, *deposits*). The practices also vary widely regarding the use of a combination of multiple subcategories (e.g., *volume/value of e-money transactions*, *by type of transaction* and *by geographic location*, and *by type of distribution channel*).

➤ Periodicity of DFS reporting varies, too, which might be related to the intensity of and the approach to supervision, but also to different preferences across authorities.

➤ More extensive financial data on [nonbank e-money issuers](#) are collected more commonly in advanced economies than in most EMDE, where the focus is on operational data. This may reflect the varying approaches to supervision of nonbank e-money issuers.

➤ In some jurisdictions (e.g., Austria, Brazil, Hong Kong, Singapore) nonbank e-money issuers that do not reach a minimum threshold can operate without a license but they may be subject to simplified reporting. In others (e.g., the United Kingdom), smaller nonbank e-money issuers need to apply for a license, but they also are subject to lighter reporting requirements.

➤ The DFS reporting requirements in many EMDE have inconsistencies, duplications, and gaps. The following are some examples:
  ■ Banks subject to detailed reporting on agent banking while nonbank e-money issuers are subject to much more limited reporting on their agents. When a single agent regulation applies to banks and nonbanks, the reporting framework is homogenous (e.g., in Brazil, Colombia, Ghana, Mexico, Paraguay, Peru).

(Continued)

- Unclear or inconsistent use of key terms within and across report templates (e.g., e-money issuer templates versus electronic retail payment statistics templates), such as *transactions*, *e-money*, *mobile money*, *agents*, *channel*, etc.
- Significant inconsistencies and gaps on consumer complaints reporting requirements. In many countries, there is no single regulation on complaints handling mechanisms applicable to all operations of banks and nonbanks. Consumer protection requirements, including reporting of complaints, are embedded in specialized regulation such as agent banking regulations and regulation on e-money issuers. This results in a fragmented framework with conflicting reporting requirements.
- Duplication of data required by different departments or authorities with conflicting periodicity or format for the same data. These problems could be eliminated or significantly reduced through better coordination. The following are some examples:
  - A mobile money regulation requires reporting monthly transaction volumes and values while a regulation on electronic channels requires quarterly volumes and values of e-transactions that include mobile money transactions
  - An authority requires granular data on e-money transactions, while another authority requires only aggregated data (high-level indicators) on e-money transactions
  - An e-money issuer is required to send the same report to two different departments of the same authority
- Duplicate data and unnecessary data fields within the same report template (e.g., when a report requires total volume and value of transactions, there is no need to require *average size of transaction*). Unnecessary data fields can lead to higher compliance costs and increase the risk of inaccurate reported data.
- Gaps such as requiring breakdown of number of e-money customers by account level (know-your-customer [KYC] tier), but failing to require the same breakdown for total outstanding value of e-money.

> In at least two countries where nonbank e-money issuers are regulated and have been operating for several years, there is no regular collection of data for supervisory purposes. Data are collected on an ad-hoc basis only.

3.2 Financial data on nonbank e-money issuers

The majority of countries in this study collect financial statements and financial indicators from nonbank e-money issuers (commercial banks that issue e-money are subject to stricter requirements set in prudential regulations).[10] Financial data are most often collected annually within the sample countries, although some authorities collect them monthly or quarterly. A few EMDE do not require reporting of any financial statement (balance sheet, cash flow, income statement). Among the report templates for nonbank e-money issuers that are publicly available or have been shared by authorities, the most extensive is the one used by the Commission de Surveillance du Secteur Financier (CSSF) in Luxembourg, which is described in Table 4.

---

10 The stricter financial reporting requirements on banks may also apply to limited-purpose banks that operate as e-money issuers, such as payments banks in India and niche payments banks in Mexico.

TABLE 4. Financial reporting by nonbank e-money issuers in Luxembourg

Item	Type of information	Periodicity
1	Balance sheet (see CGAP [2017])	Monthly
2	Identification of third-party funds held and incorporated in the balance sheet	Monthly
3	Statement of authorized reinvestments (see template)	Quarterly
4	Capital adequacy (see template)	Quarterly
5	Profit and loss account (see template)	Quarterly
6	Analysis of shareholding (see template)	Yearly

CSSF focuses on prudential control,[11] such as safeguarding funds collected by e-money issuers (items 1–3) and the sufficiency of capital (items 4 and 5). Although prudential control is probably equally important in other countries in this study, not all require frequent reporting in such detail (see Table 5).

In addition to financial statements and indicators, other types of financial data are often required. Nearly all countries require reporting of the total outstanding e-money issued, more commonly on a quarterly or monthly basis. Only a few countries require daily reporting (e.g., Bangladesh, Myanmar, Peru). Only a few countries require specific reporting of the fees generated by the e-money operations (i.e., paid by clients), and at least a couple require disclosure of the fees gained by the e-money issuer for distributing and servicing credit and insurance products for third-party lenders and insurers (see Box 3 in Section 3.3).

Most EMDE regulations require nonbank e-money issuers to safeguard the outstanding e-money by depositing an equal amount in a separate account, such as a trust/escrow account in a bank (some countries allow other types of prudentially regulated institutions to hold the trust/escrow account). Most EMDE require the e-money issuer to report the balance in this account, in a few cases daily or weekly (e.g., Bangladesh, Myanmar, Nigeria, Peru). In a few countries e-money issuers (e.g., Bangladesh, Ghana, Myanmar,[12] Nigeria, Tanzania) report on the interest gained on the account. Periodicity varies from daily to yearly. In Ghana, e-money issuers also report monthly on the distribution of the interest to customers.

3.3 Operational data on nonbank e-money issuers

E-money issuers submit more operational data more frequently in EMDE than in advanced economies among the sample countries. Operational data include data on the number and types of accounts and clients (Table 6), transaction data (Table 7), and data on e-money agents (Table 8).[13] Operational data also include information on risks and other

---

11 CSSF is not responsible for payments oversight, which is a function of the central bank. The central bank collects retail payment statistics that include e-money transaction volumes and values. This is a common division of labor in countries where financial supervision is not the responsibility of the central bank. Retail payments statistics collected by central banks are described in Section 3.5.

12 In Myanmar, this seems to apply only to institutions that are regulated by the 2016 Mobile Financial Services Regulation. Providers operating bank-led schemes under the 2013 Directive on Mobile Banking do not seem to be subject to this requirement.

13 The tables in this section may not reflect recent changes to reporting requirements in the researched countries. For example, at least three sub-Saharan African countries in the sample are migrating toward high-frequency granular e-money transaction data collection, i.e., reporting of every single e-money transaction.

TABLE 5. Example of financial data requirements for e-money issuers

Categories and subcategories	Periodicity				
	Yearly	Quarterly	Monthly	Weekly	Daily
Financial statements and indicators					
Balance sheet					
Income statement		*	**		
Capital adequacy		*	*		
Fee revenues		*	*		**
Outstanding value of e-money issued					
Total outstanding e-money issued				*	
Average outstanding in the period		**	*		
Peak outstanding during reporting period*			*		
Total outstanding by geographic location*			*		
Total outstanding by account level (KYC tier)*		*	*		
Total outstanding by type of client (legal/natural person)	*	*			
Total value in inactive accounts		*			
Total value in dormant accounts		*	**		
Fund safeguarding					
Trust account balance, by bank holding the trust				*	
Interest earned in the trust account	*			*	*
Authorized investments	*		*	*	

Note: Darker color means most common periodicities. No color means no country has that periodicity for that requirement.
*Just one country has the requirement.
**Only two countries have the requirement.

qualitative information, which are described in sections 3.5 and 3.6.

Regulatory requirements to report identification information on e-money clients were found only in Uruguay, where the central bank is to maintain a register of e-money users,[14] and in Namibia, where business e-money account holders (i.e., entities holding e-money accounts that are used for bulk payments such as salaries) are subject to a detailed report to the central bank about the beneficiaries of those payments.[15]

Although the research identified the use of a wide range of subcategories to describe types of e-money transactions, most countries stick to just a few subcategories (e.g., *withdrawals*, *deposits*, *transfers*, *payments*), which are reported mostly on a monthly basis.

---

14 Circular 2198/2014, Central Bank of Uruguay, which changed art. 88 of the Compilation of Payment System Regulations. This research did not look into how or whether this requirement is being implemented in practice.

15 Guidelines for Issuers of Electronic Money & Other Payment Instruments in Namibia, March 2012, Item 4.2.

> **BOX 3. Data on digital credit and insurance distributed by nonbank e-money issuers**
>
> A few of the countries studied require nonbank e-money issuers to report on credit and insurance that are distributed through the e-money platform. This type of reporting includes the following:
>
> - On digital credit: loan applications received, number and value of loan disbursements, total number and value of loan repayments, a list identifying all partner lender, fees collected from partner lenders.
>
> - On digital insurance: number and value of insurance policies sold through the e-money platform, number and value of insurance claims paid through e-money issuers, list identifying all partner insurers.
>
> The purpose of collecting these data, from the e-money supervisor's standpoint, is not clear, although data on fees received by the e-money issuer for distributing credit and insurance products may be relevant. It is also not clear whether this reporting duplicates the data collected from the lender/insurer by the respective prudential supervisors.

Reporting of a large number of transaction types (e.g., types of payments transactions into, for instance, bill payments and merchant payments, and types of transfers into peer-to-peer [P2P] and government-to-people [G2P]) were found mostly in sub-Saharan Africa. Only two countries require reporting of volumes/values of e-money transactions broken down by several transaction

TABLE 6. Example of data on number and types of e-money accounts and clients

Categories and subcategories	Periodicity			
	Yearly	Quarterly	Monthly	Daily
Number of e-money clients				
Total registered	*	*		
Total active	*			
Total by type (legal/natural person)	*	*		
Total by gender			*	
Number of e-money accounts				
Total registered	**			*
Total opened in the period				
Total active		*		
Total inactive		*	*	
Total dormant		*	**	
Total by gender			*	
Total by account level (KYC tier)				

Note: Darker color means most common periodicities. No color means no country has that periodicity for that requirement.
*Just one country has the requirement.
**Only two countries have the requirement.

TABLE 7. Example of e-money transaction data collection

Categories and subcategories	Periodicity				
	Yearly	Half-yearly	Quarterly	Monthly	Daily
Transaction value ($) and volume (#)					
Total only (no breakdowns)					
Total by transaction type		*			
Total by account type (customer/agents, legal/natural person)					*
Total by transaction type and geographic location				**	
Total by transaction type and type of channel (agent, ATM, POS, mobile, internet)	*	*	*		
Examples of transaction types for reporting of transaction value ($) and volume (#)					
Cash in (total only)			**		*
Deposit—government-to-person			*		
Deposit—incoming international remittance					
Payments (total only)					
Payment—utilities					
Payment—loan repayment				**	
Payment—airtime top-up					*
Transfers (total only)			**		*
Transfer—peer-to-peer					
Transfer—person-to-business or person-to-government					
Transfer—business-to-person					
Transfer—bank to e-money account and vice-versa				**	
Over-the-counter					*
Cash out/withdrawal (total only)					*
Cash out—loan disbursement, insurance claim, other				**	
Nonfinancial operations (enquiries, loan applications, account opening, etc.)			*		

Note: Darker color means most common periodicities. No color means no country has that periodicity for that requirement.
*Just one country has the requirement.
**Only two countries have the requirement.

types and by location. Interviews indicated that reporting institutions may face difficulties in compiling and reporting data when several subcategories are combined. Finally, daily reporting of granular transaction data (i.e., the listing of all e-money transactions), by transaction type, was found to be the case in a few EMDE (Bangladesh, Nigeria, Rwanda, Tanzania).

In several countries, e-money issuers have established partnerships with lenders and insurers to distribute digital credit and insurance through e-money platforms. Although digital credit and insurance fit the broad definition of DFS (see Box 1), this research focused on data reported by e-money issuers about their core e-money business. Because they usually are underwritten by a third party, digital credit and insurance are not part of the regulatory reporting requirements applicable to e-money issuers in most cases. However, at least two countries in this study require the e-money issuer to report on insurance and credit products. See Box 4.

3.4 Data on agents

The reporting requirements related to using agents were overall simpler and more consistent across banks and nonbanks in countries where a single agent regulation applies to both, or where separate regulations have similar requirements. In such countries, the amount of data required is relatively smaller than in countries with separate agent regulations for banks and nonbanks. The data concentrate on numbers of agents, broken down by some subcategories, with no details of transactions per type of agents or location, and only a few combinations of subcategories (e.g., total number of agents by type of agent

and by geographic location/service delivered).

In countries where there are different agent regulations for banks and non-banks, banks are subject to stricter reporting requirements for agent banking,[16] compared to the requirements on nonbank e-money issuers. The main difference is the level of detail required for transactions that are conducted through agent banking, with volumes and values broken down by several subcategories of transaction types. E-money transactions conducted through agents of nonbank e-money issuers are reported separately, as illustrated in Table 7. Duplicate reporting requirements were identified when there were two sets of agent regulations, for example, when banks were required to report on transactions through agents to the bank supervision department and through nonbranch channels (including agents) to the payments oversight department. Coordination between departments could eliminate duplication.

Pakistan is the only country in this study that requires reporting of the number of agents in agent banking arrangements and agent transactions by type of transaction, broken down by whether the agent is exclusive or nonexclusive, and whether the agent is equipped with biometric devices (e.g., mobile phones that can read a customer's fingerprint).

Many countries require numbers of agents to be reported by geographic location (e.g., GPS coordinates, full address, or basic information on the municipality or region), while reporting volume/value of transactions by the location of the agent is more common for bank agents. One country (Uganda) requires monthly

---

16 Agent banking regulations were not found in any advanced economy and were lacking in several EMDE, including some countries where there is regulatory permission for nonbank e-money issuers to use agents.

TABLE 8. Example of data on bank and nonbank agents

Categories and subcategories	Periodicity		
	Yearly or half-yearly	Quarterly	Monthly
Agents of nonbank e-money issuers			
Total number of agents (no breakdown)	**	**	
Total number of agents by subcategories			
by geographic location	*	*	
by type of service delivered			**
by geographic location and type of service delivered			**
by geographic location and type of agent			
by type—merchant			
by type—super agent			**
by type—subagent			**
by type—agent			
by type—type of commercial establishment			
Total added/terminated in reporting period		*	**
Total number active			
Total number inactive			
Total number of agent points/outlets (no breakdown)			
Total number of agent points/outlets by geographic location			
Transactions through e-money agents (no breakdown)			
Agents of banks			
Total number of agents (no breakdown)	*	**	
Total number of agents by subcategories			
by geographic location	*	*	
by type of service delivered			**
by geographic location and service delivered			**
by geographic location and type of agent			**
by type—merchant			**
by type—super-agent			*
by type—subagent			
by type—agent			
by type—type of commercial establishment			*
Total number added/terminated in reporting period			**
Total number of agent points/outlets (no breakdown)		**	

(Continued)

TABLE 8. Example of data on agents (Continued)

Categories and subcategories	Periodicity		
	Yearly or half-yearly	Quarterly	Monthly
Total number of agent points/outlets by geographic location		*	
Total value and volume of transactions through bank agents (no breakdown)			**
Total value and volume of transactions by subcategories			
by transaction types		**	
by each agent		*	**
by agent point/outlet			*
by account level (KYC tier)			**
by geographic location of the agent			
Average transaction size (no breakdown)			*

Note: darker color means most common periodicities. No color means no country has that periodicity for the requirement.
*Just one country has the requirement.
**Only two countries have the requirement.

reporting of the number of bank branches that are responsible for agents[17] and the number of agents assigned to each branch. Uganda also seems to be the only country that requires total volume/value of agent transactions without a breakdown by transaction types, but separately requires total value (only) of deposits and withdrawals and total volume (only) of transfers/payments. At least four EMDE (El Salvador, Ghana, Russia, Uganda) require reporting of the number of (bank or e-money) accounts opened through agents.

At least 10 countries require periodic reporting of a complete list of agents (the identification of the individuals or companies who act as agents) and all agent points/outlets (e.g., outlets of a supermarket chain) where the agent provides the services on behalf of the nonbank e-money issuer or bank. The periodicity and the types of identification information required vary. Such listings may also require identification of agent network managers (super agents, etc.) and agent status (e.g., active, inactive, dormant). These agent listings are required in addition to the periodic reporting in Table 8, and notifications are required on an as-needed basis, such as when the contract with an agent is terminated. In at least four countries, the information about individual agents is not reported through periodic returns, but through an electronic agent registry (e.g., Brazil, Mexico's pensions regulator [Consar], Nigeria [see Box 4], Pakistan, United Kingdom's FCA). In at least two cases, the agent registry contains a picture of individual agents (Mexico's Consar and Bangladesh Bank).[18]

---

17 The agent banking regulation, issued in August 2017, indicates that each agent used by a bank should be assigned to a bank branch, which will then be responsible for the operations of that agent.

18 See Dias, Staschen, and Noor (2015).

BOX 4. The Central Bank of Nigeria's agent banking registry

The Payments Department of the Central Bank of Nigeria (CBN) implemented an agent banking database in 2015 to provide details of all approved agent banking relationships and agents of licensed mobile money operators. The database generates and assigns a unique code to each agent and has details on the agent's location, identification, status (active, inactive), and monthly transaction volumes and values. Both banks and nonbanks have challenged the collection of such detailed data, and CBN is working to address these challenges. The intent is to have all banks, nonbanks, and superagents[a] that report regularly to the registry to create an agent risk rating system and to allow for automatically created analytical reports for the supervisor.

a. Nigeria has a separate regulation on super agents, which are entities licensed by CBN to manage networks of agents on behalf of banks or nonbanks.

Source: Interview with the Central Bank of Nigeria

3.5 Operational risk indicators and statistics on suspicious transactions

Almost all the countries in this study require, from e-money issuers, statistics on suspicious transaction reports (STRs), which are sent to the financial intelligence unit, and statistics and descriptions of incidences of frauds, data security breaches and service disruptions, transaction reversals, etc. Only a few jurisdictions require reporting of volumes and values of failed and pending transactions (e.g., Brazil, EU,[19] France, Luxembourg, Mexico, Singapore, Uganda). In various countries, there is similar reporting of operational risk indicators specifically with respect to transactions conducted through agent banking, when there is a separate agent banking regulation. Most countries do not impose—or at least do not publish—templates for reporting operational risk indicators and STR statistics. (See CGAP [2017] for examples of reporting requirements and a report template for suspicious transactions.)

3.6 Other qualitative information and third-party data

Some of the supervisors in this study require e-money issuers to send additional qualitative information on a regular or ad-hoc basis. In France, e-money issuers submit an annual report on internal controls, and in Indonesia, nonbank e-money issuers need to present an annual business plan and an IT audit report every three years. In Luxembourg, e-money issuers are required to provide the following on an annual basis:

■ Report on financial position, including external auditor's opinion, explanatory notes, and management report

■ Proposed allocation of results (e.g., distribution of dividends or increase of reserves)

■ Internal auditor's report

■ Management report on internal controls

■ Annual report of the compliance officer

---

19 The EU regulation is the E-Money Directive (Directive 2009/110/EC). Within the EU, the relevant country regulations are the transpositions of the EU E-Money Directive and there are some differences in reporting requirements at the country level.

Other qualitative information collected on an ad-hoc basis from e-money issuers in the researched countries include business plans, financial projections, policies and procedures, operation manuals, outsourcing agreements (e.g., with agent network managers or super agents), board meeting minutes, contingency and continuity plans, marketing materials, trust account agreements, explanations about resolution of consumer complaints, etc.

Supervisors may also use third-party data in addition to information reported by regulated institutions. This is less common in the researched countries with regard to DFS, but there were a few reported cases of media monitoring by the DFS supervisor to spot potential misconduct and of information sharing by other authorities, such as the telecommunications regulator. Third-party surveys, such as financial inclusion surveys and the results of mystery shopping, were used by a couple of interviewed authorities to identify market risks and trends and monitor financial inclusion indicators.

The downside of using the data described in this section is that a large part of it is unstructured. As a result, gathering and analyzing these data are usually labor-intensive. Technology solutions, such as SupTech, can help supervisors manage these challenges (see Section 7).

3.7 Electronic retail payment statistics

A wealth of data is collected on electronic payment transactions and payment transaction points/devices (e.g., ATM, POS, agents) across bank and nonbank regulated institutions in the researched countries. These are often referred to as *electronic retail payment statistics* and are commonly collected by payments oversight departments/units at central banks (see CGAP [2017] for an example of reporting requirements). There is less variation in these data across the researched countries compared with the e-money and agent data discussed in the previous sections.[20] In most researched countries, electronic retail payments statistics are collected separately from and in addition to data collected by the DFS supervisor (when they are different), but there is often an overlap between the two types of reporting. There is often a lack of harmonization of the definitions and terms used by these two types of reporting.

Electronic retail payment statistics comprise high-level indicators (i.e., aggregated data) on the following:

- *Number of transaction points/devices* (e.g., ATM, POS, agents, payments terminals, sometimes broken down by geographic location and available functions [e.g., whether ATMs allow deposits])

- *Number of transaction accounts* (e.g., credit card, debit card, e-money, prepaid cards, internet and mobile banking [i.e., subscriptions to internet/mobile banking service])

- *Value and volume of transactions*

  - *By a few types of transactions and payments instrument* (e.g., deposits, withdrawals, enquiries, transfers, merchant purchases by instrument [e-money, electronic fund transfers, cards])

  - *By type of channel and payment instrument* (e.g., POS [agent], POS [merchant], ATM, internet banking, e-commerce, mobile phone-based transaction at agents)

---

20 This greater homogeneity could be related, in part, to the standardized periodic reporting of payments statistics to international bodies such as the Committee on Payments and Market Infrastructures (CPMI), the World Bank, and the International Monetary Fund. CPMI publishes the Red Book for its 24 member countries, and there are several similar publications that include nonmembers, such as the Green Book on the Southern Africa Development Community, the Yellow Book on Latin America and the Caribbean, EMEAP's Red Book on East Asia and the Pacific, the Silver Book on the Commonwealth of Independent States, and the Blue Book on Europe. See https://www.bis.org/cpmi/paysysinfo.htm.

4. STANDARDIZATION: A CRUCIAL ELEMENT IN DFS DATA COLLECTION

4.1 Report templates, dictionaries, and taxonomy

Standardized regulatory data are needed to enable comparisons over time and across reporting institutions. Standardization helps to ensure data accuracy and reduce the amount of time the supervisor spends with data interpretation, cleaning, correction, and related post-submission iterations with reporting institutions. Standardization includes all or most of the following:

- Common reporting format

- Common content (scope) and frequency

- Common definition for each data field (and its data points) in a report template

- Common formula for each data field that requires calculation

Only a couple of researched countries collect nonstandardized DFS data, that is, data are submitted in a format freely chosen by each reporting institution. It is safe to affirm that most supervisors globally (and all but one interviewed supervisor) use report templates as the principal means to standardize data collection. See CGAP (2017) for examples of DFS report templates.

Report templates include data fields. Some data fields have more than one data point (e.g., the field *average e-money account balance* comprises *total outstanding e-money issued*, which is to be divided by *total number of e-money accounts*). Some data fields that are not

carefully designed and defined in consultation with reporting institutions can result in inconsistencies and errors, because different institutions often use similar terms to mean different things, as suggested by the interviews conducted for this research. A common data dictionary (and related taxonomy) are needed, although most of the EMDE supervisors in this study have not created these documents.[21] Some DFS concepts used across the countries and that need to be standardized include the following:

- **Agents.** It is not always clear how the following words in a report template are defined: *agent*, *agent point*, *agent till*, *transaction device*, *cash point*, *super agent*, *master agent*, *merchant*. These words can mean different things for different institutions.

- **Mobile wallet, e-money/mobile money account**. These terms can mean different things across country authorities and across different institutions within a country.

- **Location.** For example, the location of a transaction can be an agent's declared address, the agent/client GPS location at registration, the agent/client GPS location at the moment of transacting, or the client's registered address.

- **Transaction types.** Sorting transaction categories (e.g., cash-in, cash-out, transfers) into subcategories requires careful standardization because these terms can have different meanings across reporting institutions. When standardization is not

---

21 XBRL is a well-known taxonomy in financial regulatory reporting (not specific to DFS). It is used, for instance, by several U.S. financial regulators, the European Banking Authority (EBA), the Superintendence of Banks, Insurance and Pension Administrators in Peru, and many others for reporting financial data by regulated financial institutions such as banks (see www.xbrl.org). Although XBRL was not the focus of this research, some interviewees suggest that XBRL is of limited relevance to DFS data.

done well, there will be a high incidence of *other transactions* in the reported data, as suggested by some of the interviewed authorities.

- **Client/account types.** For example, in DFS it is useful to differentiate transaction volumes and values into retail (final) customer/accounts and wholesale (corporate) customers. The latter includes agent and merchant accounts, which may be subject to higher transaction limits in certain country regulations.

- **Completed, pending, failed transactions.** These categories are important for monitoring operational risk; however, they usually mean different things for different institutions.

- **Active, inactive, dormant accounts and agents.** Most countries in the study consider an account or agent active when there is at least one transaction in the past 30 or 90 days, but not all country regulations define these terms. Dormant agents or accounts are usually those that have been inactive for a certain period.

A dictionary/taxonomy may be especially important for young and fast-growing industries such as DFS for financial inclusion in EMDE because of the lack of common industry standards and fast-evolving business practices.[22] In addition, a few interviewees mentioned that working out a dictionary/taxonomy in detail in coordination with reporting institutions can reveal imprecisions in regulations and can help the parties solve such imprecisions.

4.2 Additional guidance to reporting institutions

Many DFS supervisors provide additional guidance for reporting institutions, such as instructions or methodological notes that may cover other areas of regulatory reporting not limited to DFS. This may include compiling regulatory requirements and interpreting key provisions, creating filing rules (e.g., periodicity, reporting channels/tools), validating rules and formulas, and so forth. The level of detail of guidance varies widely across researched countries. See CGAP (2017) for examples.

---

22 Although not related to DFS, EBA's Single Data Point Model, which focuses on bank prudential data reporting, may provide ideas for EMDE supervisors that are developing guidance to improve DFS data collection (see CGAP [2017] and http://www.eba.europa.eu/risk-analysis-and-data/reporting-frameworks).

5. CONSIDERATIONS ON GRANULAR DFS DATA

5.1 Understanding aggregated and granular data

One key element of regulatory data collection is determining the appropriate level of data granularity. Every data field in a report template (e.g., *total outstanding e-money issued*) can be traced to the underlying raw (and granular) data in an institution's IT systems that were created during the institution's operations.

Institutions use different programming languages, formats, and taxonomies for their own internal purposes, which leads most supervisors to impose standardization (*translation*), so that the reported data can be compared across reporting institutions. In addition to standardizing the data, most supervisors also require the data to be transformed— aggregated—from its original granular form into indicators such as *total outstanding e-money issued*.

For the institution, reporting aggregated data entails transforming and aggregating raw, granular data—a process that can be automated or done manually or a combination of both. Multiple data fields in a report template or in different templates can be calculated using the exact same underlying granular data. Given these two aspects, why are aggregated data rather than granular data collected?

5.2 Advantages and disadvantages of granular data

Collecting granular data is a widely discussed topic in advanced economies, mostly regarding bank prudential data.[23] The discussion revolves around the stricter prudential regulatory standards introduced in the wake of the 2008 global financial crisis, which require supervisors to collect larger amounts of more detailed data.

A major advantage of granular data is greater flexibility. Supervisors themselves calculate any desired indicator and make sure that the calculations are correct and are the same across reporting institutions. They can generate any report in any given day and conduct a limitless number of analyses by manipulating, combining, and finding relationships across data points because supervisors are not constrained by predefined indicators. Most of today's EMDE DFS supervisors are not accustomed to working with granular data because they have been relying on aggregated data that are collected through report templates. If adequate analytical skills and resources are available, then granular data have the potential to provide richer supervisory insights. Depending on the data collection mechanism, granular data also could reduce compliance costs for institutions and increase data quality (see sections 6 and 7).

The disadvantages of collecting and using granular data include having to manage a great deal of data, which could require increased storage capacity, different analytical tools at the supervisory authority, and higher transfer speeds to receive the data.[24] Also, granular data collected using traditional report templates can become very cumbersome

---

23 Relevant institutions involved in discussions about data granularity for financial sector supervision include standard-setting bodies such as the Financial Stability Board (www.fsb.org) and supporting organizations such as the Irving Fisher Committee on Central Bank Statistics at the International Finance Corporation (www.bis.org/ifc), in association with the International Statistical Institute (www.isi-web.org).

24 Storage capacity can be expanded by using cloud computing, when adequate Internet connection and speed exist. Financial supervisors may need to take extra measures to mitigate the risk of storing granular regulatory data in third-party equipment and accessing it through the Internet.

for institutions to manage. In Mexico, a large retail bank can take more than two days just to transfer granular loan data to both the central bank and to the bank supervisor (in different formats). A few EMDE in this study require reporting of all e-money and banking transactions conducted at agents through report templates, and they face problems with data quality.

Even if report templates are not used, collecting and using granular data does not eliminate the risk of errors, because raw data still need to be transformed into a standard format that is defined by the supervisor. Also, more advanced analytical and programming skills may be needed to create detailed validation rules and taxonomies to guide reporting institutions, which in turn require a more profound understanding of regulated businesses by the supervisor as compared to collecting aggregated data.

Having more data is not always better or needed. Collecting a large quantity of data without clear objectives or analytical capacity may hurt the supervisor's credibility. Several supervisors in the study noted that collecting excessive data without identifying clear uses could trigger accusations of abuse of power. In addition, some supervisors have limited legal powers to gather data. Also, reporting institutions may be more comfortable with sharing only predefined and limited high-level indicators, rather than granular data, and may resist a shift to granular data reporting. Finally, by potentially changing public expectations about the supervisor's role (from assessing risk management practices to avoiding failures or micromanaging institutions), collecting granular data could garner criticisms against the supervisor should an institution fail or face major problems.

5.3 Granular DFS data in the researched countries

The vast majority of countries in this study collect only aggregated DFS data, with variations on the breakdowns into subcategories, as described in Section 3. However, at least part of the DFS data in some of the countries are granular, as indicated in Table 9.

There seems to be a growing interest among EMDE supervisors in granular DFS data. It has not been determined that granular data are needed for effective DFS supervision, but supervisors who want to use granular data should ask some key questions:

- What extra insights could granular and periodic data on DFS transactions provide compared to the aggregated data currently collected?

- How would granularity impact compliance costs at reporting institutions, especially those catering to previously excluded clients and small institutions?

- Is there analytical capacity to standardize, validate, and analyze granular data?

- Would the supervisor be risking its reputation or credibility? Is there any legal impediment for collecting granular data?

- Would a reform of the data collection mechanism be required to avoid excessive compliance costs and ensure data quality?

Emerging technologies for regulatory data collection could help (see Section 7). BNR, for instance, has recently reformed its data collection system to shift from aggregated to granular data, including DFS data (see Box 5).

TABLE 9. Granular DFS data in interviewed countries

Country	Authority	Types of data
Bangladesh	Bangladesh Bank	Monthly reporting of all mobile financial services transactions above Tk 20,000 (approx. US$250), by transaction type, showing the name of the account holder, the amount transferred, and the recipient of the funds.
Mexico	Consar	Near real-time access to individual transactions conducted by pension administrators at the IT system of the switch used by pension administrators, with information on the pension agent who conducted the transaction (transactions include account opening, transfer of balances across pension administrators, withdrawals, deposits, etc.), GPS location of the agent at the time of the transaction, client information, and other details.
Mexico	CNBV	Monthly reporting of the number, type, and value of transactions conducted at each agent, with agent details (e.g., location).
Nigeria	CBN	Daily reporting of all individual mobile money transactions by each mobile money operator.
Pakistan	SBP	Monthly reporting of the number, type, and value of branchless banking transactions per each agent.[a]
Rwanda	BNR	Near real-time access to individual mobile money transactions at the IT systems of mobile money operators.
Tanzania	TCRA	Near real-time access to individual mobile money transactions at the IT systems of mobile money operators.

a. In Pakistan, *branchless banking transaction* means a transaction conducted at agents of bank institutions, often through mobile phones, and electronic transactions between branchless banking accounts, according to the Branchless Banking Guidelines.

BOX 5. National Bank of Rwanda's granular data collection system

The National Bank of Rwanda (BNR) started collecting disaggregated data from regulated institutions in 2017. It uses a completely automated data collection mechanism. In addition to core financial data from banks and other regulated institutions, the system includes some DFS data from mobile money operators and money transfer operators, and retail payment statistics. All data are stored in a central data repository at BNR. There is an ongoing effort to improve standardization of data points, beginning with prudential bank data. Granular operational and financial data are pulled automatically by BNR systems from the institutions' IT systems, using a virtual private network channel and end-to-end encryption. Data on consumer complaints and frauds are uploaded manually for the time being.

The reform in the data collection system is intended to reduce time and costs, reduce delays in reporting, and increase scope, quality, and reliability of data.

———

Source: Interview with BNR and Kamali and Randall (2017).

6. DATA COLLECTION MECHANISMS RELEVANT TO DFS

Some DFS supervisors are leading efforts to modernize the data collection mechanism in their organizations. A data collection mechanism comprises various systems and processes put in place by the supervisor and the reporting institutions to handle the data. The reporting to the supervisor includes systems and procedures to generate, compile, transform, validate, and send data. At the supervisory authority, it includes the systems and procedures to receive, validate, store, and share data. In broad terms, data collection mechanisms can be (*i*) paper-based copies that are sent to the supervisor and (*ii*) electronic, including the following:

- **Separate files** (Excel, PDF, Word, TXT, etc.) that are sent electronically (email attachments, web portal, or another file transfer system) or provided on hard media (e.g., memory stick). Reporting institutions can use completely automated processes to gather and validate the data in the required format before sending or (more often) use a mix of automated and manual processes.

- **Input approach,** in which sets of standardized granular data are uploaded automatically by the institution into a database at or accessed by the supervisor.

- **Pull approach,** in which raw granular data are extracted (pulled) by the supervisor from the institution's IT systems.

6.1 Automated versus manual processes

"Processes which limit the need for manual operations are considered to be best practice" (EBA 2017). All the supervisors interviewed concur that the less human intervention in data reporting/collection the better. Errors and data losses are likely to happen if someone on the reporting or collecting end manually selects data, imports/exports data, enters data into a spreadsheet, calculates indicators, etc. However, nearly all institutions and supervisors interviewed use a combination of automated and manual processes or reporting and collecting.

Manual processes persist for a variety of reasons (see some examples in Box 6), but the reasons generally relate to how data are created and managed by the IT systems at reporting institutions and how difficult it is for each institution to transform its raw data into the format required by the supervisor. A digital credit provider in the study mentioned that the IT systems were simply not made to generate standardized regulatory data. A mobile money provider said it would rather have the supervisor collect raw data directly from its IT systems, so that its staff would not need to spend time filling in confusing templates that keep changing.

Even if an institution has the raw data needed to fill report templates, system adaptations are often needed to standardize and aggregate the data. The difficulty and the cost involved in adapting systems may be more significant in large institutions with numerous legacy systems that may be incompatible with each other or with the supervisor's reporting system and thus require multiple layers of compilation before the data can be reported.[25]

---

25 BCBS (2015) notes that less than half of the largest global financial institutions have the data architecture and IT infrastructure needed to automate data aggregation and reporting. IIF (2016) observes that IT challenges and regulatory impediments often make aggregation of risk data a manual, labor-intensive task. This finding is corroborated by the U.S. Federal Reserve (2015). Another study in East Asia finds that most banks take information from legacy platforms and manipulate that data manually in spreadsheets, in what can be a costly process that increases the chances of error (Md Nur 2017). See also Dwyer (2015) and Piechocki and Dabringhausen (2016).

BOX 6. Examples of why institutions use manual procedures for reporting data to supervisors

Reasons for using manual processes to report regulatory data include the following:

- Multiple systems (e.g., product systems or modules, AML/CFT, customer care, core banking, mobile money, etc.).

- Systems that are incompatible with the system used by the compliance unit to fill in the report templates.

- Reporting formats that create difficulties given the IT architecture of an institution.

- Systems housed at different departments or outside the institution (e.g., AML/CFT, customer care).

- Gaps or errors in the automated aggregation and standardization of raw granular data.

Difficulties also may be caused by the system put in place by the supervisory authority. For instance, the recently created web-based data reporting portal of one of the interviewed supervisors requires mobile money providers to fill in 11 templates (each template is presented as an Excel worksheet), each with numerous indicators that use similar or the same underlying data, in a time-consuming process that involves manual procedures.[26]

For some mobile money providers, the system used to manage the compliance function and generate regulatory reports is not equipped to automatically access other systems that have the required data, such as anti-money laundering and combatting the financing of terrorism (AML/CFT) and customer complaints systems. Compliance officers end up receiving and handling data aggregated by other units. Virtually any obstacle in data aggregation and transformation could be solved with IT solutions, but this involves money and time, which are not always readily available for a supporting function such as compliance.

Although some supervisors perceive that imposing automated reporting processes on small institutions would be too costly, interviews with other supervisors suggest otherwise. Small and less complex institutions with a lean staff structure may be able to move more easily toward, and benefit from, a fully automated process (assuming they are computerized in the first place). None of the interviewees suggested that non-bank entrants to the DFS market are less prepared or willing to comply with regulatory reporting requirements.[27]

6.2 Quality assurance and automated validation

It is crucial to ensure a minimum level of quality of DFS data. All the countries in the study reported that they had a validation process at both the reporting and the collecting ends, with variation on

---

26 Problems with portals for regulatory reporting (not limited to DFS data) are described by the Office of Financial Research (2015, pp. 72 and 74).

27 This assumption is sometimes raised in the literature, such as AFI (2013).

BOX 7. Examples of validation checks done by supervisors on DFS data

Automated (or manual) checks may include the following:

- Check receipt of all reports, modules in a report, by due date

- Check for data completeness (e.g., all fields filled out)

- Check for correctness: compliance with pre-established validation rules, including formulas and relationships among data fields

- Consistency checks:

  - Check for consistency and plausibility within the same report, for example, checking for an outlier indicator higher/lower than a preset value (e.g., the average of the peer group)

  - Check for consistency and plausibility compared with previous reporting periods and cross-checks with other reports in the same period (e.g., a monthly growth of revenues over x% of the average in the past three months, consistency of indicators using the same underlying data in different reports)

  - Conduct a horizontal analysis (certain elements of a report across multiple institutions) to spot consistency and quality issues

Documenting the validation process and related rules, including by describing the IT tools used and the standard procedural steps (e.g., to resolve data quality issues) can also help DFS supervisors improve data quality.

_____

Source: Interviews with supervisors and EBA (2017).

---

the scope and level of automation. Some supervisors were not fully informed on how institutions validate their DFS data before submitting the data. Our research showed that institutions use a mix of automated and manual validation procedures that are not always based on common rules defined by the supervisor, but rather on each institution's own interpretation of what checks need to be done and how.

Supervisory authorities conduct validation checks after receiving the reports, with varying levels of sophistication. In the most basic check, the supervisor refuses to accept a report that contains blank mandatory fields. However, there are many other checks for completeness and consistency (see Box 7). The systems used in advanced economies and a few EMDE in the research seem to allow for more complete and complex automated checks than what is done by most interviewed EMDE supervisors (regarding DFS data only). One possible reason for this may be that good validation can be demanding in terms of resources.[28]

Several supervisors reported that they have automated over 90 percent of validation procedures, including for DFS data, but many other EMDE lag. Many do not mandate common technical formats/

_____

28 See European Central Bank, World Bank, and Banco Central do Brasil (2012), p. 38.

languages (e.g., XBRL, XML) for DFS data reporting and many receive reports in a combination of Excel spreadsheets or even PDF files, sent by email or through an online portal. The whole process of exporting data, copying data to a common database, and running validation checks can be manual. Collaborating with reporting institutions to define the rules for validation and allow for testing periods also were identified as useful practices.

Not all supervisors can automatically control submissions for regulatory reporting (e.g., keep track of submission status and send reminders, orders for re-submission, and notifications of receipt/submission status). Manual controls can be time-consuming and can lead to low-quality or delayed data. An example of a tool to manage submissions used by the Central Bank of Brazil is described in Box 8.

BOX 8. Tools for managing reporting submissions

Managing regulatory submissions is less time-consuming and more effective when using IT tools. This is considered good practice, for instance, in Europe, where most supervisors keep automated controls of regulatory submissions (not limited to DFS data). The EMDE in this study indicated that controls of submissions (not limited to DFS) are mostly done manually using Excel spreadsheets that are not accessible outside of the department responsible for the particular report template. Automating submission management involves developing IT systems to collect, store, and periodically update master data.

An example of this approach is found in the Central Bank of Brazil, which has recently implemented an automated system—Document Submission Control (CRD)—to monitor regulatory reports due and received, respective due dates, re-submissions, error messages, etc. CRD includes DFS data and runs basic validation checks. Data curators (i.e., staff who are accountable for each regulatory report) are responsible for data quality. After a report passes the first automated validation, it is sent to specialized systems in the user units for further validation, which is done daily. CRD's dashboard on each institution is updated daily, and will be made accessible to the respective institution.

———————————

Source: Interviews with the Central Bank of Brazil and desk research.

7. TECHNOLOGY AND COLLECTION OF DFS DATA

There is a global rethinking of financial supervision that goes beyond DFS supervision. A major shift may be imminent toward a data-driven and forward-looking approach that is based on better and more data. RegTech and SupTech have the potential to fundamentally transform the scope, procedures, and techniques of financial supervision.[29] Examples of this emerging supervisory approach include real-time supervision that moves away from using historical data to accessing data in real time at the institutions' IT systems, dynamic compliance,[30] the use of application programming interfaces (APIs) by supervisors, and algorithmic regulation.[31] Machine learning is another potentially powerful tool that could help supervisors mitigate challenges in data collection and analytics.[32]

Emerging supervisory approaches require improvement of regulatory data, such as by (*i*) shifting from aggregated to granular data, (*ii*) increasing the scope of data, (*iii*) reducing the time to report, and (*iv*) reducing manual procedures. Although this global rethinking of financial supervision has focused on bank prudential supervision, it may have spillover effects on DFS supervision in EMDE.

Technology can also help expand and automate the collection and analysis of unstructured data, as is being currently explored by FCA in the United Kingdom. The manual collection and analysis of unstructured data (e.g., audit reports, consumer agreements, marketing materials, policies and procedures manuals, and internal documents [e.g., licensing files and inspection reports]) are an important element of DFS supervision and can be very time-consuming and can drain supervisory resources.

7.1 How technology may change DFS data collection

7.1.1 The traditional template-based approach

To date, the most common approach to collect regulatory data (including DFS data) is to use standard report templates in formats such as Excel, TXT, XBRL, etc. The focus is on documents or reports, rather than on the data that are needed to fill in data fields in the templates. Downsides of the template-based approach include the following:

■ Limited or no flexibility for the supervisor to create new indicators to support differentiated analyses.

■ High costs of adding granular data or widening the scope of data.[33]

■ High costs and complexity of changing reporting requirements.

---

29 See Arner, Barberi, and Buckley (2016); TPWG (2017); and Matthan (2016).

30 There is no widely accepted definition for *dynamic compliance*. According to Dwyer (2015), *dynamic compliance* means automating regulatory compliance review, by embedding rules into IT protocols to eliminate chances of noncompliance and continuously assessing whether the data reported are accurate and relevant.

31 Algorithmic regulation could determine what algorithms (e.g., algorithms governing trading platforms) should look like, allow supervisors to access the algorithms in the institutions' IT systems, and potentially empower them to change algorithms. Potential uses of algorithmic regulation and supervision in DFS include algorithm-based credit scoring for digital credit and robo-advisers.

32 Artificial intelligence has been evolving from logic-based algorithms to a data-based approach to machine learning (i.e., computers learn from identifying patterns in vast quantities of data). This technology could help supervisors analyze DFS data, and even predict the behavior of different institutions. See Bauguess (2017); BIS (2015); and BearingPoint (2012).

33 IIF (2016) notes that including a new layer of granularity into an automated report template—not specific to DFS—after it has already been implemented takes at least 12 months.

- Inconsistency of indicators across different reports even when the underlying data are the same, because of different rules used for aggregation.

- Errors in calculations and differences in interpretation of data fields.

- Missing data fields.

- High numbers of report templates being used and often with overlapping scope increase burden to reporting institutions and the supervisor.

Also, because most DFS report templates require aggregated data, the downsides of aggregating data discussed in sections 5.2 and 6.1 are also relevant.

7.1.2 Alternative tech-based approaches

The higher the desired level of granularity, frequency, and scope, the greater the need to consider reforms in the data collection mechanism. An increasing number of RegTech and SupTech providers offer solutions for regulatory data reporting and collection (not limited to DFS data). New mechanisms based on advanced technology aim to address weaknesses of the template-based approach, for example, by simplifying and reducing costs of changing reporting requirements. A key aspect of the new solutions is the shift away from a focus on documents/reports, toward a focus on the underlying data needed. New approaches include the following:

- **Data-input approach.** Each institution uses automated processes that are enabled by rules embedded into its systems to transform raw data into packages of standardized granular data according to the supervisor's specifications, and sends them to a central data warehouse. This approach does not require filling in templates. Examples include the Austrian central bank, which uses this approach for prudential data reported by banks (not for DFS data) into a third-party company (see Box 9), and Mexico's Consar, which collects standardized granular data stored at a third-party company.

- **Data-pull approach.** Raw data are sourced directly from the institutions' IT systems by the supervisor. Strictly speaking, this would not be considered "reporting" because the institution has a relatively passive role, and the supervisor could collect the data at will or on preset dates, through automated processes. Examples include Tanzania's mobile money monitoring system managed by TCRA (and accessed by BoT), and Rwanda's new reporting system, which includes DFS data (see Box 5). While the literature and some interviews indicate that this approach is used in various countries for bank prudential data collection, a few supervisors in the study raised concerns with legal implications of pulling, rather than receiving, data. Also, in this approach the supervisor is responsible for transforming raw data into standard formats and for aggregating data into relevant indicators. Finally, institutions may be weary of giving remote access to supervisors because this would allow supervisors to see what is in their systems, potentially including sensitive data that may not be directly relevant for supervision.

- **Regulatory reporting utilities.** This is a new concept where technology enables the creation of centralized regulatory reporting utilities. This would be a common interface between supervisors and reporting institutions that can provide several

> ## BOX 9. Austria's cube-based model of regulatory reporting
>
> Austria's central bank, OeNB, in collaboration with the banking industry, transformed its approach to data collection to increase granularity, timeliness, quality, and scope of data, while reducing duplication, inconsistencies, and costs (in the long run) for both reporting institutions and supervisory authorities. It nearly eliminated report templates and integrated collection across departments and authorities.
>
> In the new system (an example of "an input approach"), granular data are gathered automatically at banks' systems and sent to AuRep, a company owned by banks. AuRep runs a common software platform that works as a utility and interface between banks and OeNB (costs are shared among banks). The data at AuRep are represented by a "basic cube" that cannot be accessed by OeNB. Rather, a series of enhanced datasets, such as "loan cube" or "deposit cube," are automatically prepared following standard definitions to be accessed and used anytime by OeNB. Changes in required data do not necessitate changing any template. The only thing needed is a single coordinated update to all banks, which is done through AuRep. Ad hoc needs by OeNB or the Financial Markets Authority can be mostly satisfied using the granular data in the enhanced cubes. Despite the huge amount of data being collected and the initial investments needed to build AuRep's infrastructure, a reduction of reporting costs compared to the previous template-based mechanism is expected.
>
> A presentation by Oenb on the new data collection system is found in CGAP (2017).
>
> _____
>
> Source: OeNB interviews and Piechocki and Dabringhousen (2016).

services to supervisors and reporting institutions, including a compilation of regulatory requirements and related dictionary/taxonomy translated into computer codes (i.e., instructions that can be read by computers) instead of the need for a staff person at reporting institutions to read the rules and convert them into programming codes to trigger automated processes for regulatory reporting. The utility would also function as a repository of data, including for different authorities that need such data. This concept could fully automate the regulatory reporting process, create efficiencies, and cut costs.[34]

New approaches to data collection may mean major paradigm shifts for the DFS supervisor, but several of the supervisors interviewed for this study have implemented or are planning to implement reforms to their regulatory data collection mechanism by using SupTech (e.g., Austria, Ghana, Malaysia, Nigeria, Pakistan, Peru, the Philippines, Rwanda, Tanzania, Singapore). Many reforms mostly cover non-DFS data, but they could cover any type of data, including DFS financial and operational data.

One more aspect to consider is that some RegTech and SupTech solutions rely on technologies that may not be well-understood, well-accepted, or legally

_____

34 A presentation on a concept reporting utility (not focused on DFS) designed for the U.K. financial system during a RegTech event organized by FCA is available at https://www.youtube.com/watch?v=BqieH_Bm2-U.

permitted in certain countries. For instance, cloud computing is commonly used in RegTech and SupTech solutions to lessen the costs of increasing storage capacity at individual supervisory authorities. However, cloud computing may encounter regulatory barriers in some countries.

7.2 Improving DFS data collection as part of a wider strategy

DFS supervisors in some of the EMDE in the study are leading efforts to modernize data collection in their organizations. However, some authorities lack an overarching strategy and effective interdepartmental coordination, which may result in fragmented, duplicate data collection efforts. As a result, a state-of-the-art mechanism for DFS data collection may be created, while an outdated one is used for bank financial data. The attention DFS data in EMDE is receiving could be an opportunity to push for encompassing reforms of data collection mechanisms to support effective supervision. This makes sense because DFS supervision is not done in isolation of bank supervision, or market conduct supervision, and at least part of the data collected for DFS supervision is relevant for other supervisory areas, and vice-versa.

Another important consideration is that improving DFS or any data collection will bring limited benefit if there is no adequate analytical capacity at supervisory authorities to transform the data into supervisory intelligence.[35] Supervisory authorities in the study's advanced countries and some EMDE reported having highly skilled data and supervisory analysts who are capable of processing large sets of granular data and making complex analyses using multiple analytical applications to generate timely intelligence. Unfortunately, this is not the case in all countries, and many EMDE where DFS for financial inclusion are booming may lack either (or both) the specialized human resources or the right software. Capacity may be lacking to map data needs in the first place, and to standardize data effectively.

Finally, it might be useful to look at a country, as a whole, to assess whether improvements are needed and whether synergies can be exploited across authorities involved in collecting and using DFS data. For instance, efforts should be made to find out whether data collected by one authority are useful for the DFS supervisor and vice-versa, and whether there is duplication of DFS data collection. As mentioned, in Tanzania the central bank is coordinating with TCRA to resolve the duplication in mobile money transaction reporting and to make the data collected by TCRA useful for DFS supervision by the central bank. The efforts in Austria to integrate the bank prudential data needs of different authorities into a single data collection effort by the central bank are also a useful example for DFS supervisors in EMDE.

---

35 This aspect has been noted in AFI (2013).

8. CONCLUSION

"Technology advances rapidly, giving rise to opportunities and challenges."

Jon Nicolaisen,
Deputy Governor of Norges Bank,
at the Financial Industry's Digital Services Conference,
Oslo, 9 June 2016

DFS is part of a global makeover in financial services and have been instrumental in expanding the frontier of financial inclusion in EMDE. The relevance of DFS puts pressure on EMDE supervisors to improve DFS data for supervisory purposes.

Good regulatory data, including DFS data, may be difficult to obtain in a lot of situations, but this is particularly so for supervisors in EMDE, who are more likely to lack adequate data, capacity, and expertise, as well as the financial resources and political support to improve them. However, because EMDE lack heavy legacy systems, practices, infrastructure, and fully developed regulatory and supervisory frameworks that hold them back, their supervisors may be able to more easily adopt RegTech and SupTech solutions and leapfrog into the future. The gradual steps taken by, and the approaches currently adopted in, advanced economies may not be appropriate for EMDE. Rather, better data collection mechanisms that have embedded flexibility to adapt to fast changes made possible by RegTech and SupTech can help DFS supervisors in EMDE address many challenges. Report templates, which are so deeply rooted in the supervisory mindset, and their related shortcomings, may be a thing of the past.

The main takeaways from this research can be summarized as follows:

- DFS reporting requirements (broad categories) do not vary substantially across countries in this study, but there is wide variation in the subcategories of DFS data, particularly transaction types.

- Several problems such as inconsistency in the use of DFS concepts, duplicity of reporting requirements, and other weaknesses that may lead to poor data quality or higher regulatory compliance costs exist in many EMDE.

- Only a few EMDE supervisors have created comprehensive guidance for DFS reporting, such as data dictionaries and taxonomy, which are important for ensuring data quality.

- Fast growth of RegTech and SupTech creates opportunities for EMDE supervisors to transform DFS data and related supervisory approaches.

- DFS data collection should be part of broader, authority-wide efforts to improve data quality and reduce the risks of excessive, redundant, or insufficient data collection.

- DFS reporting requirements should be based on a detailed mapping of data needs. It is important to collaborate across departments in the supervisory authority and across relevant authorities to reduce duplicate data collection and harmonize DFS concepts.

- Quality DFS data require close collaboration with reporting institutions, especially to define dictionaries/taxonomy and to implement new requirements or changes in the data collection mechanism.

- Granular data have advantages and disadvantages but it is not clear whether granular data are needed for effective DFS supervision. Since the data collection mechanism has a great influence on the quality and cost of collecting granular data, requiring granular DFS data may require a reform of the data collection mechanism.

- DFS supervisors may find it useful to have a data governance policy or strategy to drive the process of continuously improving regulatory data.

Finally, to prepare for data-driven, real-time DFS supervision, DFS supervisors in EMDE need to invest in their data analytics capacity, not only in their data collection mechanisms. Stronger data analytics could help EMDE supervisors become more confident in dealing with DFS and can support healthy and sustainable financial inclusion in the long run. Without it, improving data collection will have limited impact. The current attention given to DFS could be a great opportunity for EMDE supervisors to secure resources to improve their analytical capacity and leapfrog into an era of data-driven supervision.

ANNEX 1. RESEARCHED COUNTRIES

Jurisdiction	Type	Desk research	Phone interviews and/or questionnaires	Country visit
Afghanistan	EMDE	✓		
Austria	AE	✓	✓	✓
Bangladesh	EMDE	✓		✓
Bhutan	EMDE	✓		
Brazil	EMDE	✓	✓	
Cambodia	EMDE	✓		
Colombia	EMDE	✓	✓	
El Salvador	EMDE	✓		
European Union	AE	✓		
France	AE	✓	✓	
Ghana	EMDE	✓		✓
Guatemala	EMDE	✓		
Hong Kong	AE	✓		✓
India	EMDE	✓	✓	
Indonesia	EMDE	✓		
Jordan	EMDE	✓		✓
Kenya	EMDE	✓		
Luxembourg	AE	✓	✓	
Malaysia	EMDE	✓		✓
Malawi	EMDE	✓		
Mexico	EMDE	✓	✓	✓
Myanmar	EMDE	✓		✓
Mozambique	EMDE	✓		
Namibia	EMDE	✓		
Nigeria	EMDE	✓	✓	
Nicaragua	EMDE	✓		
Pakistan	EMDE	✓		✓
Paraguay	EMDE	✓	✓	
Peru	EMDE	✓	✓	✓
Philippines	EMDE	✓	✓	
Romania	EMDE		✓	
Russia	EMDE	✓		
Rwanda	EMDE	✓	✓	
Singapore	AE	✓		✓
Sri Lanka	EMDE	✓		
Tanzania	EMDE	✓	✓	✓
Thailand	EMDE	✓		

(Continued)

Jurisdiction	Type	Desk research	Phone interviews and/or questionnaires	Country visit
Uganda	EMDE	✓	✓	
United Kingdom	AE	✓		✓
Uruguay	AE	✓		
United States	AE	✓		

ANNEX 2. PARTICIPATING ORGANIZATIONS

Country/region	Organizations
Global	Vizor Software
	BearingPoint
	SGS
	GVG
	Alliance for Financial Inclusion
West Africa	Orange Money
Austria	Oesterreichische National bank
	Financial Markets Authority
Bangladesh	Bangladesh Bank
	Dutch Bangla Bank
	SureCash
	Grameenphone
	bKash
Brazil	Central Bank of Brazil
Colombia	Financial Superintendence
	Banca de las Oportunidades
France	French Prudential Supervision and Resolution Authority (ACPR)
Ghana	Bank of Ghana
	MTN Mobile Money
	Airtel Money
	Tigo Cash
	Vodafone Ghana Mobile Financial Services
Hong Kong	Hong Kong Monetary Authority
	Standard Chartered Bank
	Securities and Futures Commission of Hong Kong
	SuperCharger
	Monexo
	KYC Chain
	FinChat
India	Reserve Bank of India
	Iris Software
Jordan	Central Bank of Jordan
	JomoPay
Luxembourg	Commission de Surveillance du Secteur Financier
Malaysia	Bank Negara Malaysia (BNM)
	Securities Commission of Malaysia
	Propellar CrowdPlus
	MCash
	ReCal

(Continued)

Country/region	Organizations
Mexico	National Banking and Securities Commission (CBNV)
	National Pension System Commission (Consar)
	Mastercard
	Bankaool
	Mimoni
	Bancomer
	Kubo Financiero
Myanmar	Central Bank of Myanmar
	Wave Money
Nigeria	Central Bank of Nigeria
Pakistan	State Bank of Pakistan
Paraguay	Central Bank of Paraguay
Peru	Superintendence of Banks, Insurance and Pensions
Philippines	Central Bank of Philippines
Romania	National Bank of Romania
Rwanda	National Bank of Rwanda
Singapore	Monetary Authority of Singapore
	Fern Software
	NextMoney
	Fastacash
Tanzania	Bank of Tanzania
	Tanzania Communications Regulatory Authority
	Tanzania Insurance Regulatory Authority
	Vodacom Tanzania
	Jumo
	Airtel Tanzania
Uganda	Bank of Uganda
United Kingdom	Financial Conduct Authority
	ClauseMatch

ANNEX 3. BIBLIOGRAPHY

AFI (Alliance for Financial Inclusion). 2014a. "Mobile-Enabled Cross-Border Payments." Guideline Note No. 14, August. http://www.afi-global.org/publications/1598/Guideline-Note-14-Mobile-Financial-Services-Mobile-Enabled-Cross-Border-Payments

———. 2014b. "Consumer Protection in Mobile Financial Services." Guideline Note No. 13, March. http://www.afi-global.org/publications/1453/Guideline-Note-13-Consumer-Protection-in-Mobile-Financial-Services

———. 2014c. "Supervision and Oversight of Mobile Financial Services." Guideline Note No. 12, February. http://www.afi-global.org/publications/1451/Guideline-Note-12-Mobile-Financial-Services-Supervision-and-Oversight-of-MFS

———. 2013. "Mobile Financial Services: Regulatory Reporting." Guidance Note No. 3, March. http://www.afi-global.org/publications/1027/Guideline-Note-3-Mobile-Financial-Services-Regulatory-Reporting

Ainger, Neil. 2016. "Basel Banking Supervision Edict Stretches Banks' Risk Data Governance." *ComputerWeekly.com*, 5 May. http://www.computerweekly.com/feature/Basel-banking-supervision-edict-stretches-banks-risk-data-governance

Arner, D., J. Barberi, and R. Buckley. 2016. "Fintech, RegTech, and Reconceptualization of Financial Regulation." *Northwestern Journal of International Law and Business*, October. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2847806

Bank of England. 2016. "Fintech Accelerator." Press Release, 17 June. http://www.bankofengland.co.uk/Pages/fintech/default.aspx

Bates, Chris, and David Murray. 2016. "Regulation Is a High Tech Science. Is Your Firm Ready?" 7 July. http://financefeeds.com/corvil-abide-regulatory-reporting-synergy

BCBS (Basel Committee on Banking Supervision). 2003. "Risk Management Principles for Electronic Banking," July. http://www.bis.org/publ/bcbs239.pdf

———. 2010. "Principles for Enhancing Corporate Governance," October. http://www.bis.org/publ/bcbs176.pdf

———. 2012. "The Core Principles for Effective Banking Supervision," September. http://www.bis.org/publ/bcbs230.htm

———. 2013. "Principles for Effective Risk Data Aggregation and Risk Reporting," January. http://www.bis.org/publ/bcbs239.pdf

———. 2015. "Progress in Adopting the Principles for Effective Risk Data Aggregation and Risk Reporting," January. http://www.bis.org/bcbs/publ/d399.htm

———. 2016. "Report on the Impact and Accountability of Banking Supervision," July. http://www.bis.org/press/p150702.htm

BearingPoint. 2016. "Regulatory Technology: Rethinking Regulation." Service Overview. http://www.bessgmbh.com/ecomaXL/files/RegTech_Brochure_BearingPoint.pdf

BIS (Bank of International Settlements). 2016a. "Market Intelligence Gathering at Central Banks." Markets Committee. December. http://www.bis.org/publ/mktc08.htm

———. 2016b. "Financial Inclusion and the Fintech Revolution: Implications for Supervision and Oversight." Speech by Jaime Caruana, General Manager BIS, Basel, 26 October. http://www.bis.org/speeches/sp161026.pdf

———. 2016c. "Payments Aspects of Financial Inclusion." http://www.bis.org/cpmi/index.htm

———. 2015. "Central Bank's Use of and Interest in 'Big Data'." Irving Fisher Committee. http://www.bis.org/ifc/publ/ifc-report-bigdata.pdf

Burnside, R. 2000. "Towards Best Practice for Design of Electronic Data Capture Instruments." Research Paper, Statistical Clearing House. Belconnen, Australia: Australian Bureau of Statistics. http://www.ausstats.abs.gov.au/ausstats/subscriber.nsf/0/6FED60580B66EA71CA257130007D4F62/$File/1352055036_Nov2000.pdf

CGAP. 2017. "Data Collection by Supervisors of DFS: Examples, Cases, and Tem-

plates." Microfinance Gateway. https://www.microfinancegateway.org/library/data-collection-supervisors-dfs-examples-cases-and-templates

CPMI (Committee on Payments and Market Infrastructures). 2016a. "Payments Aspects of Financial Inclusion," April. http://www.bis.org/cpmi/publ/d144.htm

———. 2016b. "Fast Payments—Enhancing the Speed and Availability of Retail Payments." November. http://www.bis.org/cpmi/publ/d154.pdf

Deloitte. 2014. "The Single Supervisory Mechanism (SSM): The Big Data Issue." Deloitte and EMEA Center for Regulatory Strategy. https://www2.deloitte.com/lu/en/pages/banking-and-securities/articles/ssm-big-data-issue.html

Dwyer, John. 2015. "The Growth of RegTech: Automating Regulatory Compliance in the Investment Management Industry." *Simcorp.com*, 1 November. http://www.simcorp.com/en/insights/journal/growth-of-regtech-and-regulatory-automation

EBA (European Banking Authority). 2017. "Peer Review of the ITS on Supervisory Reporting Requirements: Final Peer Review Report." February. https://www.eba.europa.eu/documents/10180/1720738/Final+report+on+peer+review+on+ITS+on+supervisory+reporting.pdf

ECB (European Central Bank). 2002. "Task Force on Portfolio Investment Collection Systems." Final Report. Frankfurt: ECB, June. https://www.ecb.europa.eu/pub/pdf/other/portfolioinvestmenttaskforce200206en.pdf

———. 2010. "Centralized Securities Database." In Brief. Frankfurt: ECB, February https://www.ecb.europa.eu/pub/pdf/other/centralisedsecuritiesdatabase201002en.pd

Ernst & Young. 2012. "Setting the Pace of Change." 2011 Federal Reserve: Bank Regulatory Reporting Survey. http://www.ey.com/Publication/vwLUAssets/Federal_Reserve:_bank_regulatory_reporting_survey/$FILE/ErnstYoungFRBsurveyreportfinal.pdf

———. 2015. "Regulation Now: The New Standard and How Firms Are Adapting." 2015 Federal Reserve regulatory reporting survey. http://www.ey.com/Publication/vwLUAssets/EY-2015-federal-reserve-regulatory-reporting-survey/$FILE/EY-2015-federal-reserve-regulatory-reporting-survey.pdf

European Central Bank, World Bank, and Banco Central do Brasil. 2012. "A Practical Guide for Retail Payments Stocktaking." Financial Infrastructure Series. Payment Systems Policy and Research, October. http://siteresources.worldbank.org/FINANCIALSECTOR/Resources/A_practical_guide_for_retail_payments_stocktaking_version_GPW_101712.pdf

Federal Reserve Bank of Minneapolis. 2016. "Use of Regulatory Report Data in Banking Supervision." Statistical Reporting Update. Minneapolis: Federal Reserve Bank of Minneapolis, March. https://www.minneapolisfed.org/publications/banking-in-the-ninth/use-of-regulatory-report-data-in-banking-supervision

Federal Reserve Bank of New York. 2012. "Regulatory Reporting Best Practices." Presentations at workshop on regulatory reporting best practices. https://www.newyorkfed.org/medialibrary/media/newsevents/events/banking/2012/2012BestPracticesPresentation.pdf

FFIEC (Federal Financial Institution Examination Council). 2006. "Improved Business Process Through XBRL: A Use Case for Business Reporting." http://www.xbrl.org/Business/Regulators/FFIEC-White-Paper-31Jan06.pdf

FSB (Financial Stability Board). 2017. "FinTech Credit: Market Structure, Business Models and Financial Stability Implications." May. http://www.fsb.org/2017/05/fintech-credit-market-structure-business-models-and-financial-stability-implications/

Global Capital. 2016. "Banks and Fintech: Better Connected," 24 August. http://www.globalcapital.com/article/z4glc8kpgz49/banks-and-fintech-better-connected

GPFI (Global Partnership for Financial Inclusion). 2016a. *G20 High-Level Principles for Digital Financial Inclusion,* September. https://www.gpfi.org/publications/g20-high-level-principles-digital-financial-inclusion

———. 2016b. *Global Standard Setting Bodies and Financial Inclusion: The Evolving Landscape.* White Paper, March. http://www.gpfi.org/publications/global-standard-setting-bodies-and-financial-inclusion-evolving-landscape

Guardian, The. 2017. "AI Watchdog Needed to Regulate Automated Decision-Making," 27 January. https://www.theguardian.com/technology/2017/jan/27/ai-artificial-intelligence-watchdog-needed-to-prevent-discriminatory-automated-decisions).

IIF (Institute of International Finance). 2011. "Risk IT and Operations: Strengthening Capabilities." June. https://www.iif.com/publication/regulatory-report/strengthening-approaches-risk-appetite-and-risk-it-should-be-key-risk

———. 2016. "RegTech in Financial Services: Technology Solutions for Compliance and Reporting." March. https://www.iif.com/system/files/regtech_in_financial_services_-_solutions_for_compliance_and_reporting.pdf

IMF (International Monetary Fund). 2016. "Cashing In on the Digital Revolution." *Finance and Development*, Vol. 53, No. 2, June. http://www.imf.org/external/pubs/ft/fandd/2016/06/ndungu.htm

IntelligentHQ. 2015. "Are You Ready for Algorithmic Regulation?" October. https://www.intelligenthq.com/technology/are-you-ready-for-algorithmic-regulation

Kamali, Wilson, and Douglas Randall. 2017. "Leveraging SupTech for Financial Inclusion in Rwanda." World Bank blog post, 8 June. http://blogs.worldbank.org/psd/leveraging-suptech-financial-inclusion-rwanda

Matthan, Rahul. 2016. "RegTech Will Change the Way We Regulate: The Emerging Field of RegTech Promises an Algorithmic Alternative to Regulation That Can Reduce Friction and Simplify Compliance." *livemint.com*, Opinion, 26 October. http://www.livemint.com/Opinion/8WcbKsj7my7ZrPiuN9mOkL/RegTech-will-change-the-way-we-regulate.html?facet=print

Md Nur, Nurdianah. 2017. "Why RegTech Will Be Asian Banks' Next Big Focus Area." *MIS Asia*, 8 May. https://www.mis-asia.com/tech/applications/why-regtech-will-be-asian-banks-next-big-focus-area/

Middleton, Bob. 2017. "Data Governance for Self-Service Analytics." Tableau Software. White Paper. https://www.tableau.com/learn/whitepapers/data-governance-self-service-analytics

Mihaescu, Mircea. 2015. "The New Age of Algorithms." *BankNXT,* 22 April. https://banknxt.com/49972/new-age-of-algorithms/

Moody's Analytics. 2014. "Regulatory Reporting for Basel I, II, & III." Product brochure. http://www.moodysanalytics.com/~/media/Brochures/Enterprise-Risk-Solutions/Regulatory-Reporting/Regulatory-Reporting-for-Basel-I-II-and-III.pdf

Mueller, Jackson. 2016. "Enough with D.C.'s Mixed Messaging on Fintech." *American Banker*, 16 September. https://www.americanbanker.com/opinion/enough-with-dcs-mixed-messaging-on-fintech

Nejman, M., O. Cejnar, and P. Slovik. "Improving the Quality and Flexibility of Data Collection from Financial Institutions." Monetary and Statistics Department, Czech National Bank. https://www.bis.org/ifc/events/5ifcconf/nejman.pdf

Nicolaisen, J. 2016. "Digital Challenges for the Payment System". Speech by Jon Nicolaisen, Deputy Governor of Norges Bank, at the Financial Industry's Digital Services Conference, Oslo, 9 June. http://www.bis.org/review/r160613f.pdf

OFR (Office of Financial Research). 2015. "Financial Stability Report." Washington, December. https://www.financialresearch.gov/financial-stability-reports/2015-financial-stability-report/

———. 2016. "Developing Best Practices for Regulatory Data Collections." U.S. Treasury. Viewpoint 16-01, May. https://www.financialresearch.gov/viewpoint-papers/2016/05/10/developing-best-practices-for-regulatory-data-collections/

OFR (Office of Financial Research). 2017. "Promoting Higher Quality and Lower Cost in Financial Regulatory Reporting." 12 January. https://www.financialresearch.gov/from-the-director/2017/01/12/promoting-higher-quality-and-lower-cost-in-financial-regulatory-reporting/

Piechocki, M., and T. Dabringhausen, T. 2015. "Reforming Regulatory Reporting: From Templates to Cubes." BearingPoint presentation, Irving Fischer Committee on Central Bank Statistics workshop "Combining Micro and Macro Financial Statistical Data for Financial Stability Analysis: Experiences, Opportunities and Challenges." Warsaw, Poland, 14–15 December. http://www.bis.org/ifc/publ/ifcb41o.pdf

SEC (Securities and Exchange Commission). 2017. "The Role of Big Data, Machine Learning, and AI in Assessing Risks: A Regulatory Perspective." Presentation by Scott W. Bauguess, Acting Director and Acting Chief Economist, DERA, US Securities and Exchange Commission, Champagne Keynote Address at OpRisk North America, New York, 21 June. https://www.sec.gov/news/speech/bauguess-big-data-ai

Skinner, Chris. 2017. "The Semantic Regulator (#RegTech Rules)." Blog post. *The Finanser. com*. 1 February. thefinanser.com/2017/01/semantic-regulator-regtech-rules.html/

TPWG (Transatlantic Policy Working Group FinTech). 2017. "The Future of RegTech for Regulators: Adopting a Holistic Approach to a Digital Era Regulator." TPWG FinTech, June. http://new.innovatefinance.com/reports/future-regtech-regulators-adopting-holistic-approach-digital-era-regulator/

UNSWA Australia. 2016. "Exploring the Nexus of Digital Financial Services (DFS) and Financial Stability." Presentation by the Digital Financial Services Research Team, UNSW Australia, at the Alliance for Financial Inclusion conference "Maximizing the Power of Financial Access: Finding an Optimal Balance between financial Inclusion and Financial Stability." Bali. https://clmr.unsw.edu.au/resource/digital-financial-services/exploring-the-nexus-of-digital-financial-services-%28dfs%29-and-financial-stability

Viñals, J., J. Fiechter, C. Pazarbasioglu, L. Kodres, A. Narain, and M. Moretti. 2010. "Shaping the New Financial System." International Monetary Fund, IMF Staff Position Note SPN 10/15. Washington, D.C.: IMF, October. https://www.imf.org/external/pubs/ft/spn/2010/spn1015.pdf

World Bank Group. 2016. "Banking Supervisors and External Auditors: Building a Constructive Relationship: Supervisor's Insights." Centre for Financial Reporting Reform. The World Bank, Austria, and the European Union. http://siteresources.worldbank.org/EXTCENFINREPREF/Resources/survey_final_lq.pdf

Zetzsche, Dirk A., Ross P. Buckley, Janos N. Barberis, and Douglas W. Arner. 2017. "Regulating a Revolution: From Regulatory Sandboxes to Smart Regulation." European Banking Institute Working Paper Series, 2017-11. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3018534

ANNEX 4. GLOSSARY[36]

Agent registry	A database maintained by the supervisory authority in which regulated banks and nonbanks are required to provide updated information about each agent with whom the bank or nonbank works or has worked. Some country authorities use an agent registry in lieu of requiring periodic reporting of detailed agent lists and notifications of new/dismissed agents.
Aggregated versus granular data	DFS reporting requirements are usually high-level aggregated data, such as *total outstanding e-money issued* and *total value of transactions*, which are calculated through formulas defined by the supervisory authority or by the institution (e.g., total outstanding e-money issued equals the sum of all e-money client account balances at the end of the day prior to the reporting date, after all fees are deducted). Aggregated data can also be reported as an indicator, such as percent *of active e-money accounts*. Granular data are presented in a higher level of detail. There are different levels of granularity. At the most granular level, data are similar to those that are generated on an ongoing basis by financial institutions in the course of their businesses (e.g., client tables, transaction tables, loan tables, access logs, etc.). Requiring a list of all agents and all transactions per agent, or a list of all e-money transactions, is considered granular data, even though this type of reporting requires some data aggregation. Both aggregated and granular data can be reported in raw or standardized formats.
Application programming interface (API)	A set of requirements that dictate how two pieces of software talk to each other. Basically, it enables computer programs to directly communicate with one another.[a]
Audit trail (also, audit log)	A security-relevant chronological record, set of records, and/or destination and source of records that provides documentary evidence of the sequence of activities that have affected a specific system at any time, which may reflect operation, procedure, transaction, or event. A complete audit trail includes not only completed actions, but also attempted actions.[b]
Algorithm	A self-contained sequence of actions to be performed (in mathematics and computer science). Algorithms can perform calculation, data processing, and automated reasoning tasks.[c] It is increasingly being used in financial services as a substitute for decision-making.

---

36 See also BIS (2016 and 2003).

Big data	A term that describes the large volume of data—both structured and unstructured—that inundates a business on a day-to-day basis. Big data can be analyzed for insights that lead to better decisions and strategic business moves.[d] Big Data refers to data sets that are so large or complex that traditional data processing application software is inadequate to deal with them. Tools for big data analytics collect, organize, and analyze vast amounts of data to discover patterns and other useful information.[e] The term *big data* often refers simply to the use of predictive analytics, user behavior analytics, or certain other advanced data analytics methods that extract value from data, and seldom to a particular size of data set.[f]
Data	Information that has been translated into a form that is efficient for movement or processing. Data can be stored in different file formats that can be used in different machine types. Database, database management system, and relational database technologies help organize data. Data may include text, audio and video, log, and web activity records.[g]
Data aggregation	The process of compiling granular data and other information from databases or other sources with intent to prepare combined datasets for data processing or reporting. Data are searched, gathered, and presented in a report-based, summarized (aggregated) format.[h]
Database	Organized information. It is a collection of information organized in such a way that a computer program can quickly select desired pieces of data. It is an electronic filing system. Traditional databases are organized by fields, records, and files. Most databases contain multiple tables that may each include several fields. For example, an institution database may include tables for products, employees, and financial records.[i] The power of the database is its ability to make data comprehensive, so that they yield useful information. A database query language, such as standard query language (SQL), allows a database administrator to interact with the database.[j]
Data category and subcategory	High-level fields in reporting templates that may be broken down into several other data fields called subcategories. An example of DFS data category is *total volume of transactions*. This category could have several subcategories, according to the types of transactions to be reported, such as *transfers*, *deposits*, *withdrawals*, etc. Each subcategory could be divided into further subcategories (e.g., G2P transfers, P2P transfers).
Data collection	The gathering of data by financial supervisory authorities. Although it can involve data requested on an ad hoc basis, this paper focuses on periodic (regular) data collection.
Data collection mechanism	The combination of the systems and procedures used to report (institution's side), and the systems and procedures used to collect, validate, store, manage, and access the data (supervisor's side).
Data field	A space to be filled out in a report template. It can comprise one or multiple data points (see *data point*).

Data governance	The overall management of the availability, usability, integrity, and security of data employed in an enterprise. A sound data governance program includes a governing body or council, a defined set of procedures, and a plan to execute those procedures.[k]
Data need	Closely related to the concept of reporting requirement. The data need is the data or information that the supervisor needs to fulfill its policy and supervisory objectives. Data needs are often (fully or partially) translated into formal requirements to the regulated institutions, which will need to report data to the supervisor and can face consequences for not complying with such requirements.
Data point	Each piece of data that is needed to fill out a data field in a report template. A data field may comprise a single data point or multiple data points put together through a formula or other type of aggregation. Data points may comprise aggregated data such as high-level indicators (e.g., total e-money issued) that are calculated through formulas run by the reporting institution (e.g., the sum of all e-money client account balances). They can also reflect reporting of highly granular data (e.g., reporting of each e-money account and all fields (attributes) that characterize each e-money client account, such as account number, opening date, balance.
Data warehouse	A relational database that is designed for query and analysis rather than for transaction processing. It usually contains historical data derived from transaction data, but it can include data from other sources.[l]
Digital financial services (DFS)	The broad range of financial services accessed through digital devices and delivered through digital channels, including payment, credit, savings, and remittances. Mobile financial services are a subset of DFS. DFS can be offered by banks and nonbanks. Digital devices can be mobile phones, cards combined, ATMs, computers connected to the internet, and any other device that connects to a digital transactional platform. Customers can transact through branches, but mostly through agents or remotely from their digital device. They typically transact on small-value transactional accounts targeted at the mass market, but the term as used in this paper also includes models where customers transact over the counter (e.g., at an agent). Despite this broad definition, this paper focuses on e-money issuers, the use of agents by banks and nonbanks, and retail payment statistics.
DFS data	Any data required by financial supervisory authorities regarding the operations of DFS and DFS providers.
DFS provider	Any type of bank or nonbank institution that offers DFS to the public.
DFS supervisor	A supervisory authority, or the team inside it, that is responsible for DFS, including e-money issuers, electronic retail payments, and digital (or electronic) channels.
Dictionary	A document in which the definitions of all fields of a regulatory report template are organized. It is closely related to *taxonomy*.
Digital credit	A type of DFS, more specifically loans, that are commercialized through digital channels, such as mobile phones. These are usually low value, use alternative credit scoring methodologies, and have fast processes.

Granular data	Data before it is compiled and aggregated into high-level indicators, such as total e-money issued. The total e-money issued by an institution is calculated and reported based on the compilation of data on the e-money account system, which has, for instance, all client accounts' balances and transactions. Not all granular data are *raw data*.
E-money	An electronic store of monetary value on a technical device that may be widely used for making payments to entities other than the e-money issuer. The device acts as a prepaid bearer instrument that does not necessarily involve bank accounts in transactions.[m]
E-money account	A prepaid instrument based on e-money that can be offered by banks and other authorized deposit-taking financial institutions, as well as by nondeposit-taking payment services providers, such as MNOs.[n]
E-money issuer	A type of DFS provider (bank or nonbank) that issues e-money.
Financial supervisor	A financial supervisory authority charged with supervising or regulating financial institutions and similar, such as a central bank, a banking commission or supervisory authority, a financial markets authority, etc. The term is used extensively to refer to authorities responsible for supervising e-money issuers and banks using agents. In the case of many countries in this study, this authority is also a central bank.
FinTech	An industry comprising companies that use new technology and innovation with available resources to compete in the marketplace of traditional financial institutions and intermediaries in the delivery of financial services. Financial technology companies consist of both startups and established financial and technology companies that are trying to replace or enhance the use of financial services of incumbent companies.[o]
Legacy systems	Potentially outdated computer systems, programming languages, or software. For banks, it often means a system that cannot be taken out of service, and the cost of designing a new system with a similar availability level is high, e.g., systems to handle customers' accounts.[p]
Master data	Descriptive information that is necessary to manage, store and forward reported data, verify completeness of submissions, and run certain data checks. It includes the business card of an institution, i.e., basic data to identify the submitter of a report and to determine the reporting requirements according to the business characteristics (e.g., characteristics of assets, conglomerate structure, etc.).
Mobile financial services	A type of DFS, primarily delivered through mobile phones. Includes mobile money, mobile insurance, mobile savings, and others.
Mobile insurance	A type of DFS, more specifically an insurance that is commercialized through mobile phones, usually using a mobile money platform. Parts of the insurance servicing (premium and claims payments) are also often done using the mobile money platform.
Mobile money	A type of DFS, more specifically a type of e-money delivered primarily through mobile phones and mobile money agents.

Nonbank e-money issuer	A type of e-money issuer that is not a bank.
Nonbank	Any regulated institution that is not treated as a deposit-taking banking institution in a particular jurisdiction.
Raw data	Data that sit in an institution's IT systems and that were created in the course of the institution's operations, using different programming languages, formats, and taxonomy for their own internal purposes. When it is created, raw data are granular. But raw data can also be and usually are aggregated for an institution's own internal purposes. Most DFS data collected are not raw, but rather standardized into common formats.
RegTech	Technology solutions to improve compliance at potentially lower costs by regulated institutions. This includes solutions for regulatory reporting.
Regulatory report	A formal return or report that the supervisory authority receives from regulated institutions, according to reporting requirements previously defined by the authority. *Regulatory report* is synonymous with *regulatory return*.
Report template	The model instrument/document created by supervisory authorities that needs to be filled out by reporting institutions. Each field in the report template is referred to as a data field. Each data field can comprise one or more data points. The report templates can have aggregate (more common) DFS data or granular (disaggregated) data.
Reporting institution	An institution regulated by a financial supervisory authority that is subject to reporting requirements.
Reporting requirement	The requirement imposed on regulated institutions to provide data to the supervisor, on a regular or ad-hoc basis, regardless of the formats or systems used for such reporting. The reporting requirements may or may not be fully described in regulations.
Return (regulatory)	Also known as *regulatory reports*.
Semantic technology and data point models	Technology that converts regulatory text into a programming language. Machine-readable regulation is an example that would allow more automation and could significantly reduce the cost of change. It could also help ensure greater consistency between the intentions of a regulation and its implementation.[q]
Standardized versus raw data	Raw (or business) data are generated by the information systems of a reporting institution. Since institutions use systems that comply with the specifications of the IT vendor and/or the institution, raw data are usually not comparable across institutions, raw data are generated at the most granular level (e.g., client tables, transaction tables), and they also can be (and usually are) aggregated at some level for internal management purposes (e.g., risk management). Standardized data are raw data that have been put into a common format that is defined by the supervisory authority, including by using common definitions and formulas. Most data collected by financial supervisory authorities are standardized, but there are examples of collection of raw data for DFS supervision (both aggregated and granular).

Structured data	Data that have been organized into a formatted repository, typically a database, so that their elements can be made addressable for more effective processing and analysis. A data structure is a kind of repository that organizes information for that purpose. In a database, for example, each field is discrete and its information can be retrieved either separately or along with data from other fields, in a variety of combinations.[r]
SupTech	Technological solutions focused on improving the processes and effectiveness of financial supervision and regulation.
Supervisor	Refers to the unit or department responsible for supervision, or to the supervisory authority, depending on the context.
Supervisory authority	Financial authorities in charge of financial supervision, such as central banks, financial market authorities, financial conduct authorities, and bank superintendence. It does not include regulators of other sectors, such as telecommunications regulators.
Suspense account	An electronic register that serves as a holding place until funds are moved to other appropriate accounts. For example, an issuer's e-money platform may place in suspense accounts the funds related to pending transfers from one customer to another, or pending merchant payments. Also known as *clearing accounts*, transitory accounts, *pending funds accounts*, etc.
Taxonomy	Closely related to *dictionary*. It determines fundamental rules for regulatory reports. For instance, for each field of a regulatory report (and included in the dictionary) the taxonomy determines attributes and interrelationships. These attributes and interrelationships reflect regulations as well as accounting rules and are the basis for programming the IT codes that automate regulatory reporting at reporting institutions.
Transformation (of data)	The translation of raw data into a common aggregation level and/or format determined by the supervisor. This process is closely related to the rules defined in the taxonomy.
Unstructured data	Data that either do not have a predefined data model or are not organized in a predefined manner. Unstructured data are typically text-heavy, but may contain data such as dates, numbers, and facts. The lack of organization and defined formats results in irregularities and ambiguities that make it difficult to understand using traditional programs as compared to data stored in fielded form in databases or annotated (semantically tagged) in documents.[s] Unstructured data is a generic label for describing data that are not contained in a database or some other type of data structure. Unstructured data can be textual or nontextual. Textual unstructured data are generated in media like email messages, PowerPoint presentations, Word documents, collaboration software and instant messages. Nontextual unstructured data are generated in media like JPEG images, MP3 audio files, and Flash video files. The information contained in unstructured data is not always easy to locate. It requires that data in both electronic and hard copy documents and other media be scanned so a search application can parse out concepts based on words used in specific contexts. This is called semantic search.[t]

Validation	The process of ensuring that data are clean and correct. It uses automated or manual routines, often based on predefined validation rules (also called *validation constraints*) that check for correctness, meaningfulness, and security of data that are input to the system.[u] Validation checks can be done at both the regulated institution (before reporting) and the supervisory authority (when and after receiving a report). This term may be used elsewhere in a more restricted manner to refer only to checks of data correctness.

a. http://www.planetcompliance.com/2017/03/16/the-regtech-glossary-a-beginners-guide-to-the-industrys-terminology/
b. https://en.wikipedia.org/wiki/Audit_trail.
c. https://en.wikipedia.org/wiki/Algorithm
d. https://www.sas.com/en_us/insights/big-data/what-is-big-data.html
e. http://www.planetcompliance.com/2017/03/16/the-regtech-glossary-a-beginners-guide-to-the-industrys-terminology/
f. https://en.wikipedia.org/wiki/Big_data.
g. http://searchdatamanagement.techtarget.com/definition/data
h. http://www.planetcompliance.com/2017/03/16/the-regtech-glossary-a-beginners-guide-to-the-industrys-terminology/
i. Webopedia.
j. http://whatis.techtarget.com/definition/structured-data.
k. http://searchdatamanagement.techtarget.com/essentialguide/Building-an-effective-data-governance-framework.
l. Oracle Help Center, https://docs.oracle.com/cd/B10500_01/server.920/a96520/concept.htm
m. https://www.ecb.europa.eu/stats/money_credit_banking/electronic_money/html/index.en.html
n. *Payment Aspects of Financial Inclusion*, 2016, Committee on Payments and Market Infrastructures and the World Bank Group, April 2016.
o. https://en.wikipedia.org/wiki/Financial_technology
p. http://www.planetcompliance.com/2017/03/16/the-regtech-glossary-a-beginners-guide-to-the-industrys-terminology/
q. http://www.planetcompliance.com/2017/03/16/the-regtech-glossary-a-beginners-guide-to-the-industrys-terminology.
r. http://whatis.techtarget.com/definition/structured-data.
s. https://en.wikipedia.org/wiki/Unstructured_data.
t. http://searchbusinessanalytics.techtarget.com/definition/unstructured-data.
u. Wikipedia.