

# Transforming data collection from the UK financial sector: a plan for 2021 and beyond

We worked with industry to develop a vision for data collection and a plan on how to get there.



## Content

### Executive summary

Background and aims

Key findings

Next steps and transformation programme

---

### Review on a page

---

#### 1: Structure of the paper

---

#### 2: Review background and activities

---

#### 3: Problems with the current reporting process

3.1: Complexity, legacy, and strategic planning

3.2: Value and collection rationale

3.3: Interpretation

3.4: Finding and sourcing data

3.5: Reconciliation and data quality

---

#### 4: Potential solution areas

4.1: Common data inputs

4.2: Modernising reporting instructions

4.3: Governance and architecture

4.4: Other solution areas

---

#### 5: Establishing a vision for data collection transformation

5.1: Common data standards

5.2: Modernised reporting instructions

5.3: An integrated approach to data collection

---

#### Box A: Learning from others

---

#### 6: Approach to data collection transformation

6.1: Delivery approach principles

6.2: Transformation programme

6.3: Alignment with other initiatives

---

#### Box B: Use case overview

Use case assessment framework

Use case roadmaps

Use case selection

---



**Annexes**



# Executive summary

---



## Background and aims

Data collection has always been an integral part of what the Bank does. Without the data we collect, we cannot identify risks, design good policy, and take action in a timely and targeted fashion.

But we, and the systems we oversee, are changing. Technological advances and automation mean that more data than ever before is being created and captured. Simultaneously participants across the financial system, including authorities like ourselves, expect more high quality, timely data to guide them in their decision making.



These changes have put strains on the current data collection process, and on the suppliers of data within the financial sector. They have increased costs, and meant systems and processes are being used in ways their designers never envisioned. But digitisation may also offer solutions to these problems. As we explored in our recent joint innovation project 'Digital Regulatory Reporting',<sup>[1]</sup> technology could transform how the data collection process works, with benefits for the Bank and industry.

In 2019, in order to decide how to respond to these changes, the Bank committed to a review of data collection. The review was launched with the publication of the discussion paper '[Transforming data collection from the UK financial sector](#)', in January 2020. The review's aim was to shape the evolution of reporting over the next 5-10 years. So its goals were long-term, and its scope was wide. It covered structured data used by our regulatory, statistical, and markets and banking teams. At the heart of its approach was open and collaborative dialogue between all parties to the data collection process. This dialogue has, so far, taken place via more than 260 internal and external events, with over 130 organisations, and through receipt and review of over 60 written responses to the discussion paper.

**This paper sets out the findings from that discussion process, and the next steps for the review.** It documents what we have learnt about the problems with data collection – and the possible remedies for those problems – before setting out our aspirational vision for data collection and how we think we will get there.

## Key findings

We think the issues the industry and the Bank face can be summarised in three key questions that need to be addressed:

- How can we ensure our data collections are worthwhile and valuable exercises for regulators and industry to invest in?
- How can industry best understand and interpret our reporting instructions so that high quality data is provided?
- How can we remove legacy data, process and technology siloes and streamline the reporting process?

By successfully addressing these three questions, we believe we can achieve our vision for data collection that: 'The Bank gets the data it needs to fulfil its mission, at the lowest possible cost to industry'. Achieving that vision means that our data collections deliver rich and rapid insights into emerging risks, and the Bank can act confidently and in a targeted way to mitigate those risks. It means that our data collection processes are lean, so we and industry can make the most of scarce resources. It means that the process and purpose of data collection is clear. And it means that data collection continues to be safe, so that data is collected and used securely, legally and ethically.

In order to realise the vision, the Bank and industry will need to transform how we manage and collect data. In particular, the review identified three key reforms that need to take place:

- **Defining and adopting common data standards** that identify and describe data in a consistent way throughout the financial sector. These common standards should be open and accessible for use by all who need them. Their adoption will bring benefits well beyond reporting.
- **Modernising reporting instructions** to improve how our reporting instructions are written, interpreted and implemented. There are a range of steps we think this will involve, from setting up better Q&A processes to potentially rewriting our instructions as code.
- **Integrating reporting** to move to a more streamlined, efficient approach to data collection. This reform includes making data collection more consistent across domains, sectors and jurisdictions, and designing each step in the data collection process with the end-to-end process in mind.

**We don't think the reforms can be achieved quickly or easily.** We are aware of the difficulties we and industry will face moving away from legacy solutions. And many of the changes required will be cultural as well as technical, with sustained investment required to make the range of improvements we have identified.

## **Next steps and transformation programme**

To help deliver these reforms we, alongside the FCA, want to set up a multi-year, and multi-phased, transformation programme. We want this programme to be delivered with an iterative and pragmatic approach. During each phase we will aim to deliver a series of 'use cases' focusing on particular collections or types of collections. Each use case delivered will





add value in its own right, as well as delivering improvements that can then be applied to other collections over time. We expect the first phase, taking place over the next 24 months, to deliver improvements and lessons for the future, but to only impact a small number of selected use cases[2]. The second phase, taking place over roughly the subsequent three years, will focus on expanding the transformation into new areas with an increased focus on integration. Subsequent phases will scale the transformation to maximise value.

And while we are aware that the global pandemic poses unique challenges to us all, we do want to make progress over the coming year. We hope by the end of this year to have the structures in place that will govern this programme of work, with delivery teams designing and testing the future of data collection.

In particular, we are looking to build a core team comprising staff from the Bank, the FCA and from the firms from whom we collect data, as well as solution providers who wish to participate. We currently estimate the team's size to be between 10-15 people, of which 12 might come from firms. We expect the core team to be supplemented by part-time input from others, both from industry and from the Bank and FCA. We will need participants who have expertise and vision to contribute to designing solutions. In addition, to ensure the new changes work for everyone, we would like individuals who are directly involved in the current data collection process to provide feedback on issues and solutions. We invite interested parties who wish to contribute resource to contact us directly by email: [DataCollectionDP@bankofengland.co.uk](mailto:DataCollectionDP@bankofengland.co.uk).



As we did during the review, we want to continue to work with industry in an open, collaborative and transparent manner. We think having the widest possible set of views will help us get to the best solutions to existing problems. And we are acutely aware we can't think about 'integrating' the reporting process without having all the parties in the reporting process involved.

Openness is also important because we do not think we can deliver these reforms via the UK transformation programme alone, and nor do we want to. We know there are ongoing and completed data collection projects taking place around the globe, some of which we are already involved in, and many of which we talked to during the review process. And, as became clear during the review, our reforms tackle fundamental issues that industry are already trying to address: in particular inconsistencies in how financial data are described and identified. So we will need to align and coordinate with these and other initiatives.

Being open and transparent is the first step in doing so, but we will also need to take an active role to align with, and in some cases possibly coordinate, broader initiatives. That is likely to be particularly important regarding international engagement. For instance, we will continue to work with other financial authorities to push the development and adoption of key financial identifiers. We will deepen our relationship with relevant private sector initiatives.

And we will engage authorities around the world to develop and share best practice, and push reporting and data standards up the global policy agenda. This international work will take time. As an initial step, we are working on a joint project<sup>[3]</sup> run by the Singapore BIS-Innovation Hub to create common data standards and machine executable reporting rules for mortgage reporting.

Together we think these programmes, and the reforms they can bring about, will build some of the soft infrastructure for the digital age. That digital infrastructure will deliver value throughout the financial system.



# Review on a page

---







## We conducted extensive research with industry



The discussion paper was discussed at >260 different external events



We spoke to >130 different organisations, including banks, insurers, regulators, service providers and trade bodies



We received over 60 written responses to the discussion paper published in January 2020

## What we learnt about problems with data collection



### Complexity, legacy and strategic planning

Multiple reporting processes within firms, legacy systems and complex reporting rules hamper change



### Value and collection rationale

Difficult to measure value or cost of data collection



### Interpretation

Difficult to use and understand current corpus of reporting instructions



### Finding and sourcing data

Data requested is sometimes not available, and can be difficult to find and extract



### Reconciliation and data quality

Difficult to reconcile similar data across reports; costly data quality controls

## Our vision

“The Bank gets the data it needs to fulfil its mission, at the lowest possible cost to industry”

## Our approach



### Long-term and vision led

Ambitious vision spanning 10 years allows time to plan, invest and sequence with existing change programmes



### Incremental and pragmatic

Tackling smaller problems that deliver incremental value, while allowing for flexibility and small fails early-on



### Open and collaborative

An open and clear approach that encourages collaboration and sharing of ideas and results with others



### Continuous value and use-case driven

Targeted delivery that creates value in a defined scope

## We think the vision needs three key reforms to be realised



Defining and adopting common data standards



Modernising reporting instructions



Integrating reporting

- Consistently identify and describe financial data
- Open and accessible by all
- Embedded in operational processes

- Improve how our reporting instructions are written, interpreted and implemented
- Range of possible options including: instructions as code and better industry engagement

- Reporting processes part of one consistent solution
- End-to-end integration of steps in data collection
- Integration of firm input data and our reports

## We intend to deliver our vision and reforms through a range of initiatives



### Establish joint transformation programme

- Work with the FCA and industry to design, test, evaluate and deliver change for data standards and reporting more widely
- Deliver value for initial use cases: 'commercial real estate database', 'quarterly derivative statistical return' and 'Liquidity Metric Monitoring Tool'



### Take an active role where appropriate

- Work with the PRA on reporting change initiatives like the Solvency II consultation
- Coordinate with other Bank areas promoting data standards at relevant international fora
- Deliver further improvements to internal systems and processes



### Align with other initiatives

- Align with private sector initiatives to standardise wholesale market processes and data
- Learn from similar reforms that took place in other industries
- Coordinate with, and learn from, other international reporting change work



# 1: Structure of the paper

---

The paper has two aims. Firstly, the paper aims to share the findings of our year-long data collection review work. That involved an intensive engagement process with internal and external stakeholders, where we sought to understand the underlying issues and assess possible solutions for improving data collection. This is the focus of the first half of the paper: chapters two, three and four. Secondly, the paper aims to set out a path towards transforming data collection, responsive both to legacy constraints and the opportunities made possible by technological innovation. We describe our vision for data collection, and our incremental, use-case based approach for delivering tangible change. We intend to start the first phase of this transformation, in close collaboration with industry, during 2021. This is the focus of the second half of the paper: chapters five to seven.

The second chapter sets out the background to the review, which was launched with the publication of a discussion paper at the beginning of 2020. The review was conducted in two phases, the first exploring the options with industry, and the second contributing to the design of the plan itself.

The third chapter discusses the current problems with data collection, weaving in the comments submitted in formal responses to the discussion as well as feedback we received during bilateral discussions with firms. It highlights that many issues start from the complexity of reporting systems, both within the Bank and within firms.

The fourth chapter summarises industry preferences for the various solution areas presented in the discussion paper: common data inputs, modernising reporting instructions, and architecture and governance. There was strong support for the first two solution areas, with respondents suggesting that the third area can only be useful once the first two are achieved.

The fifth chapter discusses the Bank's vision for transforming data collection. Our vision centres around three key reforms: common data standards, modernised reporting instructions and an integrated approach to data collection. The first two link closely with ideas previously presented in the discussion paper; the last reform outlines the need for breaking down existing data and solution siloes within firms and the Bank.

The sixth chapter describes how we plan to tackle the reforms and reach our vision. It lays out the principles we plan to use to approach this transformation, our primary vehicle for carrying making change: our joint transformation programme, and how we plan to work with other relevant initiatives and bodies to deliver the reforms.



The annexes contain additional information. Annex 1 gives more detail on how the review was carried out. Annex 2 lists the external working group members. Annex 3 discusses a stylised view of the data collection process. Annex 4 contains an overview of some of the use cases discussed within working groups.



## 2: Review background and activities

---

The review took place in response to recommendations in the [Future of Finance report](#), published in June 2019. Specifically the Bank committed to ‘Launch a review in consultation with banks, insurers and financial market infrastructures to explore a transformation of the hosting and use of regulatory data over the next decade.’ The review investigated ways to decrease the burden on industry and to increase the timeliness and effectiveness of data used by the Bank.

The Bank<sup>[4]</sup> launched its review by publishing the discussion paper on ‘[Transforming data collection from the UK financial sector](#)’ (the DP) in January 2020. The DP clarified the scope of the review to include statistical data and data collected for our markets and banking operations in addition to regulatory data. The DP also asked for feedback on the current reporting process, potential solutions for issues identified, and possible delivery approaches. The potential solutions it discussed drew on ideas explored in the [Future of Finance report](#) and the joint Bank/FCA pilot on [Digital Regulatory Reporting](#) (DRR).

The review took place throughout 2020 (see Annex 1). The external review participants were a mix of firms (who are responsible for submitting data to us for regulatory or other purposes), trade bodies (representing subsets of the financial industry), solution providers (supplying software or services to firms) and various public authorities who have an interest in our work. Collectively, we refer to all these organisations as industry. Alongside external conversations, the review included discussions and meetings with individuals from across the Bank and PRA – hereafter referred to as internal participants.

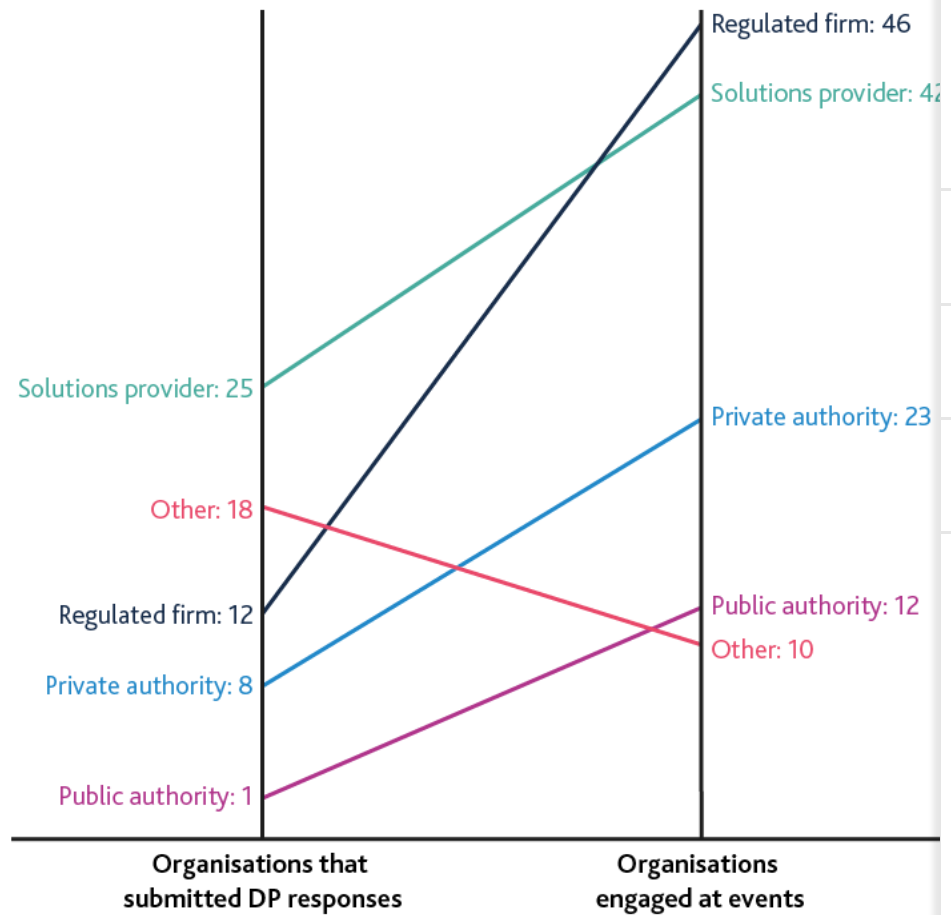
The review started with a four-month discussion period. During this phase we explored the DP questions through a process of direct stakeholder engagement, including bilateral meetings and roundtable events. We also held similar discussions internally with colleagues across the Bank. By the close of the discussion period, the Bank received over 60 written responses from industry. The evidence gathered from this period informs the first half of this report, setting out problems with the current reporting process, and summarising the feedback on possible solutions to those problems.



**Figure 1: Organisations involved in review process**

The discussion paper received **over 60 written responses** from a range of organisations

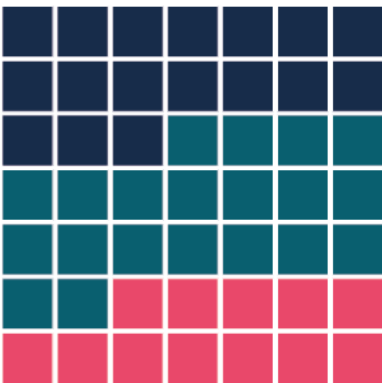
We presented or discussed the review directly with **over 130 different organisations** at over 260 events



During the second half of the review we shaped our discussion feedback into our transformation plan for data collection. During this phase we discussed the overall vision, approach, and use cases that fed directly into the plan presented within this document. The focus of this phase was a series of internal and external working group meetings. External working group members represented a range of reporting firms (retail and commercial banks; wholesale banks; financial market infrastructure; and insurance – see Annex 2). Each external working group meeting in the series of three was prefaced with an internal one, with representatives across the Bank’s business areas. For those not part of the working group process, such as technology solution providers and other standards bodies, we discussed many of the same topics in bilateral meetings and specific roundtables.

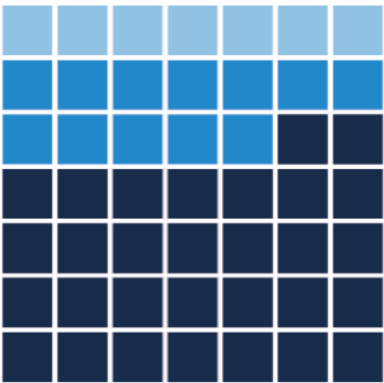
Figure 2: External working group members by organisational type[5]

All external working groups



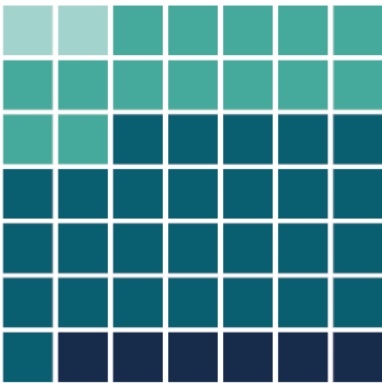
- Retail and commercial banks
- Wholesale banks and financial market infrastructure
- Insurance firms

Retail and commercial banks



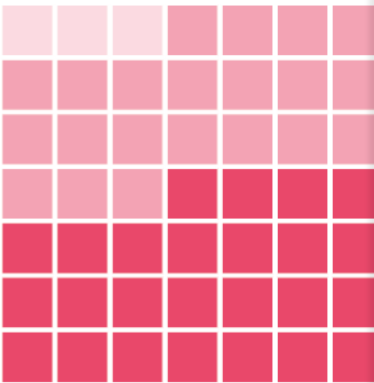
- Trade bodies
- Building societies
- UK banks

Wholesale banks and financial market infrastructure



- Central counter party
- Trade bodies
- International banks
- UK banks

Insurance firms



- Trade bodies
- General insurers
- Life insurers





## 3: Problems with the current reporting process

---



The review deepened our understanding of the current problems with data collection, and where specifically they occur. This enhanced awareness of the issues helped us frame potential solutions. In the future, it will also help us evaluate the extent to which the solutions that we implement have effectively tackled the problems identified. In addition to providing the areas of focus for this review, we hope this chapter proves useful to other parties interested in improving data collection.

Review participants confirmed many of the costs and issues we laid out in our original discussion paper, before fleshing out these problems with useful examples and ancillary points. These issues are grouped into themes and summarised in this chapter.



To frame our feedback on issues – and the solutions presented in the next chapter – we developed a stylised view of the data collection process. This process contains a number of phases which, although they should not be read as sequential, do have dependencies on each other. The reporting process outlined here could apply to either setting up a new data collection, or making a significant change to an existing one. The process for one-off collections is similar, but truncated.<sup>[6]</sup> We see the reporting process as having seven key phases:

- Analytical question and collection initiation: an authority<sup>[7]</sup> identifies a new policy question and determines what data to collect
- Rule-making governance, approval and publication: authority drafts proposals, conducts internal review, consults with firms and publishes final, approved version
- Impact analysis and interpretation: firms interpret the rules and review their impact on internal processes and resources
- Solution development and implementation: firms establish how to source, authorise and submit data to the authority
- Execution, maintenance and submission management: firms execute the designed processes, and produce the data required
- Assurance and use: authority quality assures the data submitted, and firms provide explanations or corrections as needed
- Change and review: authority reviews data collection to assess its continued relevance, and ceases collections that are no longer required

The graphic below depicts the phases, with a fuller description included in Annex 3.



**Figure 3: Rule-making, governance approval and publication**

	Authority	Firm
<p><b>1</b></p> <p>Analytical question and collection initiation</p>	<p>Prompted by a policy question, we decide to request data. We think through the practical issues related with data collection - what data, where from and so on.</p>	
<p><b>2</b></p> <p>Rule-making governance, approval and publication</p>	<p>We conduct an internal review, and take proposal through governance. We also seek feedback from firms before publishing the final version of requirements.</p>	<p>Firms submit their views on draft proposals during public consultation. Firms attempt to estimate economic and operational impact of the proposal.</p>
<p><b>3</b></p> <p>Impact analysis and interpretation</p>		<p>Firms take a deeper look at the instructions, determine what applies to them, and assess the impact on people, processes and systems (and associated costs).</p>
<p><b>4</b></p> <p>Solution development and implementation</p>	<p>We create a solution to receive, store and use the data submitted.</p>	<p>Firms translate their understanding of instructions into solutions to deliver data. The solution will need to source data, create and test reports and set-up pre-submission sign-off processes.</p>
<p><b>5</b></p> <p>Execution, maintenance and submission management</p>		<p>Firms execute their designed processes, and associated sign-offs. Firms continue to monitor changes to internal or external environments, and make necessary improvements when required.</p>
<p><b>6</b></p> <p>Assurance and use</p>	<p>We quality assure data by applying validation and plausibility checking procedures. We prepare the data for internal storage and use systems.</p>	<p>Firms may have to provide explanations for values they have reported or correct errors and resubmit.</p>
<p><b>7</b></p> <p>Change and review</p>	<p>Periodically, we review the data collections to ensure they remain relevant and cost-effective.</p>	

Review participants discussed the problems and challenges they face in the various stages of this data collection process. We hope this chapter captures many of the issues that contribute to the costs and difficulties we and industry face today. That being said, we do not

expect to be able to fix them all as part of our transformation work. Some issues relate to data collections that we do not own, for example. Others reflect constraints that are not in our power to fix.

At a high level, the review confirmed the three big cost and effectiveness challenges identified in the discussion paper. Data users need:

- timely, good quality data
- flexible data sets that can be repurposed when required
- efficient processes to collect and store data

### **3.1: Complexity, legacy, and strategic planning**

Many firms felt they face huge complexity in trying to meet our, and others', reporting requirements. For instance, one global firm said they submitted 40 consolidated reports on liquidity alone across the various jurisdictions they operate in<sup>[8]</sup>. Review participants talked about two key sources of complexity. First, complexity on the firm side: the data for reports can come from dozens of legal entities, hundreds of business lines, and thousands of systems. Even seemingly simple data points like 'total lending amount' can have over 10 definitions at some firms, and each of these may have subtle definitional differences from similar terms used by other firms in their sector. Second, complexity driven by authorities: we (and other authorities) ask for similar data (but with slightly different definitions), across multiple reports, at different breakdowns within the group (solo, group, various sub-consolidation bases).

Review participants said dealing with this complexity was one of the hardest parts of the reporting process. They felt it was both a key source of avoidable cost and that it resulted in many data quality issues. This complexity was evident during the review process itself. When large firms organised themselves in response to our review, they found it hard at times even to identify the relevant people internally to discuss issues being raised.

Complexity is perhaps highest for global firms. They are not just faced with UK reporting requirements, but similar requirements globally. They feel that they incur a high cost in navigating multiple, and often inconsistent, reporting frameworks around the world.

Some of that complexity is unavoidable. From a Bank or authority perspective, data is requested for different purposes by different business areas (e.g. prudential regulation, financial stability, statistics, markets and banking), and to meet regular, or ad-hoc needs. Differences in the uses of data result in differences in what data authorities collect and how they go about collecting it. For instance, for the Bank's collections, each legal framework that allows us to collect data<sup>[9]</sup> has its own specific steps we need to follow. It's a similar story for the data itself that we collect. The data we collect is often defined by a wider set of standards



These standards can be subtly different, causing subtly different data requirements – the accounting methods used for statistical data may differ slightly from those used in regulatory data, for example.

Similarly, some complexity at firms is also likely to remain in future. Firms emphasised that any changes to the reporting process should not impact the choice of products they offered to their customers. Larger firms said their entity structures, and business lines, were a result of complex decisions about how they wanted to run their business. They didn't expect those decisions to change going forward.

### Figure 4: Mapping problem statements to the data collection process

The problem areas we identified during the review impact different parts of the data collection process. Further, the problem areas may impact firms, an authority (like the Bank) or both. For instance, complexity is a dual firm and authority problem that affects all phases of the data collection process.



Problem Area:	Data collection process:						
	Analytical question and collection initiation	Rule-making governance, approval and publication	Impact analysis and interpretation	Solution development and implementation	Execution maintenance and submission governance	Assurance and use	Change and review
Complexity, legacy and strategic planning	■	■	■	■	■	■	■
Value and collection rationale	■	■	■				■
Interpretation		■	■	■		■	■
Finding and sourcing data			■	■	■		
Reconciliation and data quality					■	■	■

Consequences for:

■ Firm   ■ Firm and authority   ■ Authority

### Dealing with legacy

Despite this, many review participants felt that aspects of the complexity they faced today were the legacy of decisions made in the past which could be addressed. For instance, many firms' legacy systems continue to have inherently inconsistent source data that has not been



touched since the data was captured at the time of the transaction. This means the full set of data we ask for isn't always available - since legacy systems often capture a different set of data to newer systems. Or it can mean that data quality issues remain unsolved at source (see subsection on Reconciliation and data quality). Cumulatively these historic decisions have created the current complex reporting landscape. Most importantly, participants felt that these legacy issues needed to be, and could be, tackled to make real progress on reporting. From the Bank's side, we increase complexity when we create new sets of reports without fully integrating those reports into the landscape of current reports; this causes a key industry complaint that the Bank collects the same data multiple times. On the firm side, increases in complexity reflected a legacy of manual processes, siloed processes, and outdated, fragmented operational systems.



Left unaddressed, review participants said the complexity caused by legacy operational systems in particular does not just impact reporting to us. Rather it impacts the quality of a number of different activities carried out by firms. Many of the concerns raised by participants were strikingly similar to those identified in the recent report on post-trade processing produced by a [Post-Trade Technology Market Practitioner Panel](#). As the Bank's Executive Director of Markets, Andrew Hauser, put it in the foreword to that [report](#) '...complexity in post-trade matters, for three main reasons. First, it raises the cost of financial services that we all use — sometimes materially so. Second, it holds back innovation — because post-trade services provide the bedrock and data on which 'front end' services are built. And third, aged, slow or incompatible systems can pose real risks to operational resilience: an issue of great importance to firms and regulators, as we have been so vividly reminded in recent months'.

## Funding and strategic planning

Respondents to our review told us that, in part, failures to tackle these legacy issues were caused by a second problem area: a lack of resource for data collection and difficulties with strategic planning.

Inside and outside the Bank, people felt data collection was at times overly focused on meeting short-term deadlines. As we noted in the discussion paper, there has been a significant increase in data collection in recent years. Many review participants felt firms and the Bank had been making short-term decisions to cope with this change. But they felt this short-term focus prevented necessary longer-term investments. For instance, participants felt firms persist with reporting processes based on legacy systems or tactical solutions, even where there are better alternatives available.

## 3.2: Value and collection rationale

### Collection rationale

Some participants in the external working group felt that funding available within their firms to invest in reporting processes was limited, partly due to a lack of buy-in on the value of reporting. That was not because they thought reporting was a waste of time. On the contrary, firms generally thought data collection was necessary and part of good regulation. But at times, they didn't understand why we were collecting data, or have any sight of how the data was being used. Further, they felt that this lack of clarity about the rationale for specific collections had an important side effect: that it made it hard to interpret the instructions for those collections (see next sub-section).

### Cost-benefit analysis

Why we need the data drives its value to us as an organisation. During the 'rule-making governance phase' of the process, we typically estimate the value of new collections during a 'cost-benefit analysis (CBA)'. During the CBA, we check to make sure the value of a new collection exceeds its costs.

The Bank can find it hard to estimate the value of a collection. Typically, we carry out data collections to help us meet our goals<sup>[10]</sup>, such as to keep the UK financial system safe. Better data feeds into better analysis, and ultimately, better actions. But the value of any data we collect to meet those goals is hard to judge. For instance, what is the value of a firm's Common Equity Tier 1 Ratio (CET 1)<sup>[11]</sup> data point to improved financial stability? And how much more valuable is CET 1 than the total amount of a firm's mortgage lending?

All review participants felt it is hard to quantify the other side of the cost-benefit equation: cost. When asked, firms in the review process struggled to describe the costs of collections accurately, especially for new or prospective collections. They provided a variety of reasons why this was the case. They said a lot of the difficulty related to operating process and systems being used for multiple purposes. In turn this meant it is difficult to separate the cost of, say, mortgage reporting from the cost of other mortgage related activity; or the cost of supplying UK reporting from similar international reports carried out by the same system.

### Disruption and change management

One reason CBAs are such a crucial part of the reporting process is because we know changes to data collections can be disruptive to firms. External participants were keen to identify some of the ways data collection changes impacted their business.

The issues they flagged included: the high costs of setting up a project to deliver small changes; and having to pause internal projects as our mandatory changes take priority. They felt ad-hoc requests were particularly disruptive. They felt they were hard to plan for, and that



it was difficult to respond to requests in a dynamic and automated way. Global firms highlighted difficulties in managing the impact of other authorities' reporting changes on systems used to report to the Bank. Firms said they often used the same system to report similar data to multiple regulators or authorities. They said doing so was generally efficient, but changes to the system to produce third party reports also meant retesting the system for knock-on impacts to Bank reports.

### **3.3: Interpretation**

A critical early part of the process for firms, after our instructions are published, is the interpretation of reporting instructions. Reporting instructions play a critical role for all players in the reporting process. These instructions describe what needs to be reported, by whom, when and how. Nonetheless, industry participants felt that processing and understanding these instructions was one of the biggest sources of avoidable cost of the data collection process.

#### **Usability**

Many review participants from regulated firms, solution providers and RegTechs felt instructions were often not easy to use. Their problems started when they tried to obtain the full set of instructions. They said they struggled to find the latest version of the instructions, had difficulties locating all the relevant documents, and felt the website hosting instructions could be hard to navigate.

Review participants complained usability issues continued whilst they read the instructions themselves. Users of instructions complained of wasting time on dealing with bits of instructions that weren't relevant to them. This was true across industry, but the reasons why this happens differed between firms. Smaller firms tend to have simpler business models and carry out fewer activities (only deposit taking and mortgage lending for example). As a result, they felt many instructions, and parts of reports, aren't relevant to them (for instance data points applicable to derivatives related activities). Individuals working in larger firms had similar issues, but for different reasons. They were often highly specialised. They were only interested in the bits relevant to their business area, or to their role in the reporting process.

All users struggled to deal with links between instructions. Reporting instructions often contain cross-references to other documents or to other sections within the document. Users found navigating these links a slow and difficult task.

#### **Understanding**

Once they have the instructions, users need to understand what they say. For firms this typically means trying to answer two key questions: does this reporting requirement apply to us? What do we need to do to carry it out?







Industry users of instructions complained about how they are drafted. Some users felt instructions were written in over-complex legal language – they wanted a ‘plain English’ version of the instructions. Firm developers and business analysts in the review felt instructions were often overly verbose. They felt the intent of the instructions could be expressed more clearly and precisely in code or a mathematical language.

Some of industry’s difficulty in understanding is due to ambiguity in how instructions are written. As we laid out in our Discussion Paper,<sup>[12]</sup> the instructions are often deliberately limited in their level of precision and retain a degree of flexibility. This is because our requirements need to cater to a diverse and changing set of firms.<sup>[13]</sup> That being said, participants said that implementing such instructions was hard work. Firms said they sometimes needed to resolve internal disagreements between different business areas on what the correct interpretation should be.<sup>[14]</sup> To help, firms often looked to their peers for guidance on interpretation – they were keen to make judgements in line with others in industry. But they felt they lacked proper forums where those common interpretations could be agreed.<sup>[15]</sup>



### **Guidance and clarification**

To help their understanding, users of the instructions often needed the writers of the instructions to make clarifications, or provide guidance. But many users felt that when they asked, the response rates for clarifications were slow. Some smaller firms said that they were afraid to ask questions, worrying that doing so would give regulators a negative view of their firm as a whole.

For the Bank, these questions and queries often come too late. Ideally, we would like to engage with queries early in the consultation phase, before reporting instructions are finalised and implemented in our own systems. Despite extensive regulatory Q&A processes, and the support of industry trade associations, firms often only fully understand the impact of a reporting change much later, when they create and roll out solutions that carry out our instructions.

### **3.4: Finding and sourcing data**

Often queries arise where firms are not sure if they have the data; or they have the data, but are not sure how to report it. They said this can occur where the requested data does not map directly onto any operational function or system within firms. In some cases, they can deal with this problem by unpacking the logic to get from the data point requested to the data they generate and hold. In others, it means they needed to work harder to find the data we asked for.





Issues finding data were a common problem amongst review participants, particularly larger firms. Again this was partly due to complexity, with firms struggling to locate specific data points in their large, legacy estates. This was partly due to where data is stored. Data that firms regularly use for their own purposes is often stored in easy to access places, such as so called 'data warehouses'. Some data we ask for is stored in harder to find and access places, like local Excel files or in locked-down operational systems. To get hold of the data, firms said they needed to set up costly manual processes, or similarly costly new feeds and processes to change where data is stored.

A bigger problem for firms than difficulty in finding data was not having it at all, or only having the data for newer customers or products. Firms were clear that sourcing unavailable data was one of the most costly, and slow parts of setting up a new collection. There were a number of solutions firms used to fix missing data issues, each bringing costs and delays. Sometimes, the data is not available in-house, but can be sourced from third parties or requested from customers. At a minimum this requires firms to set up a new data feed, at a maximum it can require email and phone campaigns to ask for updated information. Sometimes, the data isn't generated or recorded by an internal system. This can require firms to make costly changes to one, or many, systems to meet the need.



### **3.5: Reconciliation and data quality**

The various problems that occur at each phase of the reporting process often result in generation of poor quality data. Dealing with data quality issues compounds problems that occurred elsewhere in the data collection process. Again, review participants felt complexity and legacy issues made resolving data quality problems unduly difficult.

Firms most commonly check data and apply fixes at two key places: at the input layer (the source of the data), and at the report or output level (after the transformations are applied).<sup>[16]</sup> Review participants felt this was inefficient. Ideally, they thought fixes should be applied only once, early in the process. But, they felt that doing so was often unfeasible: the direct benefits to the firm of better reporting to the Bank for that individual report were too small to warrant a costly fix in a legacy source system.

As part of the quality checking process, both Bank and external participants reconcile related data between reports. Given the size and complexity of the reporting landscape, there are sometimes thousands of data points that can, and should, reconcile. Participants found this was often hard to do. Large firms struggled to ensure the same data point tallied when it was submitted as part of multiple reports. Users of the data said at times they wondered which data to trust, after they struggled to identify the cause of differences between seemingly similar data points.

In the discussion paper, we noted how recurrent or excessive data quality issues can have a direct cost to firms in terms of fines and costly Skilled Person Reviews<sup>[17]</sup>. Given the consequences, it is understandable firms said they commit a lot of costly management time to checking and signing-off reports.



## 4: Potential solution areas

---

During the review we spent time discussing possible ways to improve data collection. Our discussions were framed by the three solution areas we laid out in the discussion paper:

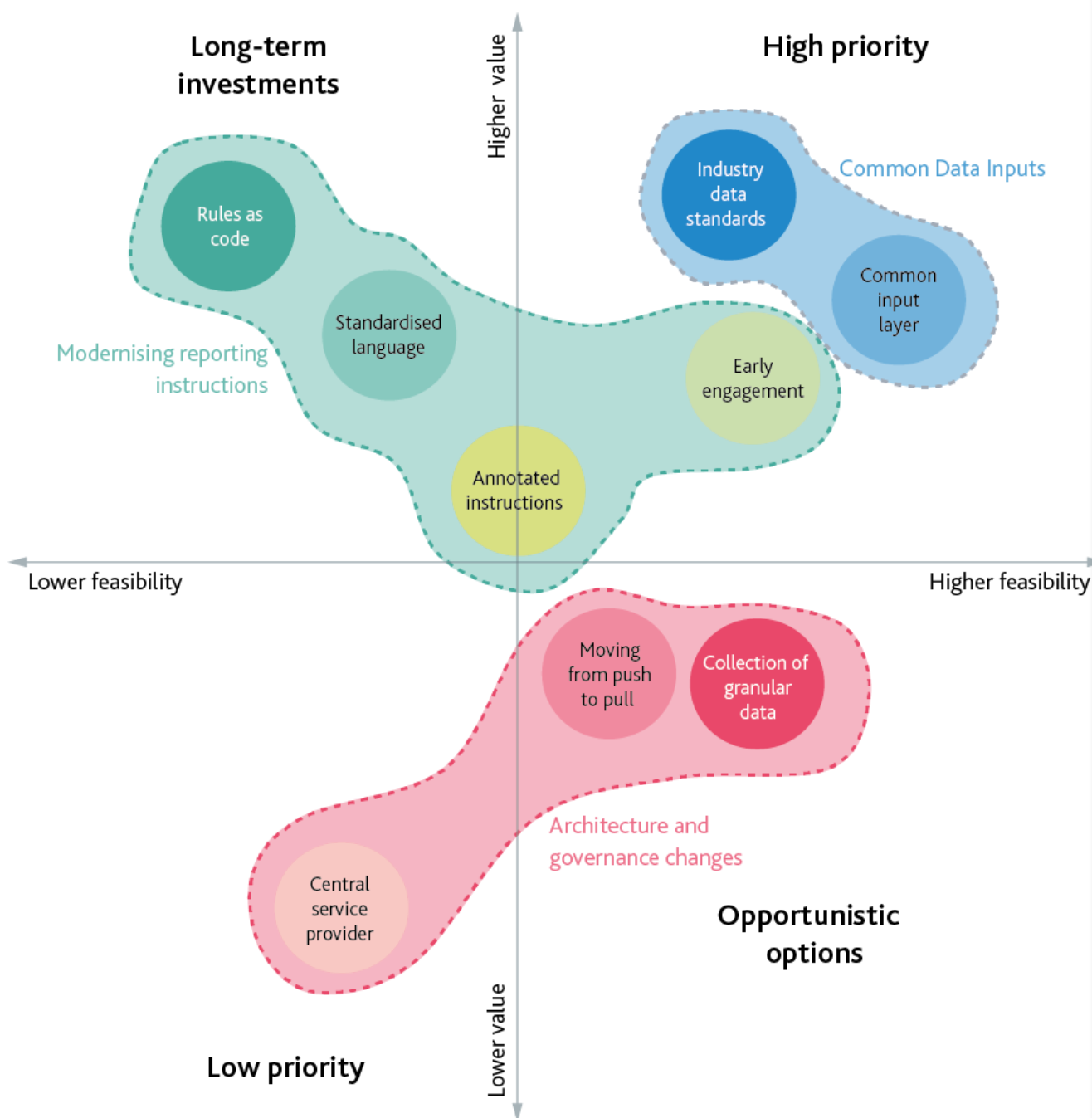
- common data inputs,
- modernising reporting instructions, and
- architecture and governance changes.

In the discussion paper, each block was then broken into more specific options.

Industry provided feedback for each option. Figure 5 visualises sentiment for each area and their options using a matrix of value and feasibility.



**Figure 5: Industry feedback on options**



The following sections summarise responses to the discussion paper questions on each block and the options that underpin them.

### 4.1: Common data inputs

Review participants generally concurred with many in the Bank that ‘common data inputs’ was the most important solution area. The options we discussed in the discussion paper were all about defining a set of underlying data inputs which could be used to rebuild a number of Bank reports. These data inputs were ‘common’ since they would be defined consistently across industry and across reports. Industry participants viewed this as the foundational element for fulfilling any vision to transform data collection. Participants felt

common data inputs could address a number of issues identified in the previous chapter: from making it easier to manage reporting complexity, to helping them find and source data. And by providing a common language to talk about data, common data inputs improve the ease and consistency of interpretation.

Participants pointed to the benefits of existing private initiatives to create input layers to make reporting easier. One global firm talked us through how producing their own internal input layer had allowed them to streamline their liquidity reporting which feeds in data from over 40 group entities. Reporting solution providers said the benefits of input layers were key to their value to their clients. They said their input layers allowed them to create multiple reports from the same set of underlying data, and protect their clients from the costs of future reporting change. Many firms that used reporting solution providers agreed with their benefits. They however felt the cost and effort of mapping their own data to the provider's common input layer was high – with set up costs accounting for up to 50% of the total cost of the solution itself.

To make this integration process easier, a common input layer could consist of a set of industry data standards used to run operational processes within firms. These standards could be reused as the basis of authority reporting. The data standards would have tighter, more prescriptive definitions than you might see in a common input layer. Some solution providers, wholesale firms and product trade bodies vocally supported this option. They felt the benefits of industry data standards would be enormous, and would allow industry to tackle the complexity and legacy issues that exist within firms. This was not a unanimous view however.

## Challenges

Some participants questioned whether standards for operational data were feasible, or even desirable. They felt that dealing with legacy systems in firms was too great a challenge and too big a risk for the rewards on offer. For instance, firms spoke eloquently about how legacy systems were about more than just IT. They noted that for some long dated financial products, like residential mortgages or interest rate derivatives, changing the system might mean changing financial payments for the product itself<sup>[18]</sup>.

Regardless of the level at which to define common data inputs, participants felt it will be hard to create them. Many participants said it would be hard to reach a fair agreement on what the common data inputs were and what they meant. For example, a smaller firm worried that large, systemic firms would dominate the process. Some firms questioned how we can achieve standardisation given the heterogeneity in firms' data, and the heterogeneity of the Bank's needs. Small firms and insurance firms were particularly vocal on this front. They questioned how we can create industry data standards for large numbers of bespoke or niche financial products.

Firms with bespoke products were amongst many that talked about the scale and complexity of the task. Larger reporting solution providers said their common input layers contained thousands of data fields and require teams of over 100 working to maintain them. For global firms, these scale issues weren't just about covering data across domains, but also ensuring the input layer covered international as well as UK based reporting. Insurers pointed out that the level of aggregation at which common data inputs make sense doesn't just depend on intended use – it is also related to how firms package and sell their financial products. An illustrative example they gave were insurance products that covered a number of risks – say, travel, home and motor insurance– which meant they were unable to break down the insurance premium paid to cover each risk individually.

### **Ways forward**

Some participants pointed to the creation of common data inputs by individual parties or sections of industry (see Box A). Moreover, they felt advances in thinking and tools meant common data inputs could be delivered at lower cost than in the past. For instance, derivatives and securities lending markets participants pointed to their work on the Common Domain Model, which standardises data and events for some of finance's most complex and bespoke products.

To make progress, what most felt was needed was strategic action from the Bank and other public sector actors. Many participants felt there was a lack of available common data inputs including industry data standards for some data domains, particularly for retail products. Internal and external participants said a lack of adoption was reducing the value of new and existing standards. They felt public authorities (including the Bank and FCA) have the powers to drive adoption of standards and should be actively doing so. For instance, one trade body talked of a recent missed chance to improve industry processes, where an authority didn't mandate the use of a product identifier in reporting for their market.





## Figure 6: Mapping data standards and modernising reporting instructions to the data collection process

Solution areas will have differing benefits for each step in the data collection process (shown below). They will also bring new costs (not shown).

Solution Area:	Data collection process:						
	Analytical question and collection initiation	Rule-making governance, approval and publication	Impact analysis and interpretation	Solution development and implementation	Execution maintenance and submission governance	Assurance and use	Change and review
Common data inputs:							
Common input layer		■	■	■	■	■	■
Industry data standards		■	■	■	■	■	■
Modernising reporting instructions:							
Annotated instructions		■	■	■		■	■
Standardised language		■	■	■		■	■
Early engagement		■	■				■
Rules as code		■	■	■	■		■



### 4.2: Modernising reporting instructions

Many participants thought common data inputs should be delivered alongside improvements in another solution area: modernising reporting instructions. There was perhaps strongest support for improving instructions among smaller firms. This may reflect differences in the problems they face relative to their larger peers. They complained less about problems finding and sourcing data within their own firms, and more about the complexity of the reporting instructions themselves.



Participants suggested that a range of low- and high-tech solutions are required. These solutions differed in their use of innovative technology, and when they can feasibly be delivered. Most of these solutions were covered in some way by our original discussion paper.

In the short-term, some firms and users viewed improving the usability of existing instructions as a priority. This meant addressing some of the issues relating to how our instructions are published and consumed (see sub-section on instruction usability issues). For instance, participants said we could invest in improving the usability of our own website where our instructions are often published. Others advocated publishing our 'instructions as data', for instance via an Application Programming Interface (API). Similar to our concept of 'annotated reporting instructions', publishing 'instructions as data' would allow third parties and machines to connect to our instructions. Our instructions could then be consumed by other third party applications, perhaps alongside instructions published by other authorities. However, some firms viewed 'instructions as data' as a low value activity, and questioned its feasibility given the complexity of some current reporting instructions.



Many external participants said that to reduce complexity, greater standardisation of instructions was required. Firms talked about how this might be achieved, from reusing terms and shrinking the size of our reporting dictionary, to trying to make our reporting instructions more consistent.

To simplify instructions, and make solution design and implementation easier, technical experts felt a mindset change was needed in how instructions are designed and structured. They advocated greater use of 'data modelling' techniques. They felt we needed to move away from creating reports, where reports are typically self-contained single tables, towards relational data, where data is located in a number of interrelated tables – building on what we do for regulatory collections that use the 'data point modelling' approach. Some saw a change in how instructions are designed as a part of a shift to 'digital first policy-making', where legal drafters, policy makers and developers collaborate on policymaking to make regulation easy to consume and implement.

Longer-term, a change in how reports are designed may pave the way for the final technical option considered: instructions as code. For some participants, instructions as code was the final destination for our reporting instructions. But there was disagreement on what 'instructions as code' meant, and therefore what benefits it might deliver. Some firms saw it as part of a process to automate the execution of reporting instructions. For them, instructions as code wasn't just about better interpretation, it was about taking humans out from large parts of the solution development and interpretation phase. Other participants felt automation wasn't important to get the benefits. They felt what was important was to express our instructions with the precision of code – some sort of pseudo-code perhaps – so humans could more easily and clearly translate the logic of instructions into executable applications.



Many external participants thought we needed more than technical improvements in how instructions are delivered. They felt industry and the Bank could talk more openly and earlier. And they felt the Bank could help industry come together to agree a standard understanding of reporting instructions. In both cases, they felt the outcome would be an easier interpretation process, with fewer queries and clarifications needed during later phases (see sub-section on instruction understanding issues and clarifications).

## Challenges

Technical experts in the review cautioned about rushing to deliver innovative, complex technical solutions at scale. This was one reason participants favoured pseudo-code over executable code. Publishing in pseudo-code would mean we wouldn't need to publish code in multiple programming languages – which would be difficult to maintain – or tie our instruction format to a particular solution or technology.

Participants also worried about any change that fully stripped out firm interpretation from the reporting process. This argument came from both sides of the reporting process. Firms were concerned that it would mean the data they submit might not be right for the question we were trying to answer – exacerbating problems with understanding the collection rationale. They felt the interpretation stage was a key step for them to check the data we asked for against its purpose of use. Within the Bank there were concerns that firms would lose responsibility for the data they were providing.

Many participants emphasised the skills and cultural changes needed to modernise reporting instructions. For instance, internal participants flagged we would likely need to boost the number of data architects and developers in our reporting teams. Many regional firms were particularly concerned, fearing the technical skills they would need locally are in short supply in their area. Though post-Covid, and given advances in remote working, this might change.

## Ways forward

As with common data inputs, participants felt modernising reporting instructions was key to transforming data collection. In general, given enough time and investment, they felt the challenges we face could be overcome. Again they pointed to a number of existing private and public initiatives that showed how we could move forward (see Box A).

### 4.3: Governance and architecture

Overall, firms were open to considering the options proposed for architecture and governance changes. However, when compared to the other solution blocks, firms considered such changes as a lower priority. Firms felt the benefits of many of the options would be small, while firms challenged the feasibility of their delivery. In particular, firms

stressed they would want confidence that the Bank and/or a central service provider are storing their data securely, correctly interpreting their data, and drawing valid conclusions about the health of their firm.

**Pull-vs current push-model.** Firms agreed that there were potential benefits of moving from a push- to a pull-model in reducing reporting costs, if it led to reducing effort in interpreting reporting instructions. But, although some participants saw the Bank ‘pulling’ data as part of the long-term vision, few felt it would deliver much benefit compared to the current ‘push’ model in the short-to medium-term.

Furthermore, many firms disagreed strongly with any suggestion that it might result in the Bank being able to pull data in real time, expressing unease about the regulator or central bank having direct access to their systems. In addition, firms had questions on the governance and security implications of a pull model, such as the mechanics of data verification, pulling and storing large volumes of data securely, and accountability in the event of a security breach.

**Central service provider.** Some firms were more receptive to a central service provider being part of a potential pull model, for instance as an intermediary between themselves and the Bank. These firms saw a model where they could push data to a central entity from which the Bank subsequently pulls data. This would allow firms to maintain data ownership, and verify their data before making it available. Subsequently, the Bank would have confidence in the quality of the data that it can request on demand, while avoiding direct access to firm systems. This operating model was familiar to firms who are required to centrally consolidate entity level returns and those that use third party solution providers as part of their reporting process.

There was no consensus however on what the optimal role and responsibility of a central service provider should be. While the functions of a central service provider could fall within the Bank’s remit, there was more support for a separate entity, possibly funded by industry, with oversight provided by the Bank. Firms also explored ways of delivering the benefits of a central service provider by the Bank approving one or many software providers. But industry participants questioned how an approved provider model would work with a competitive market of service providers.

**Increasing data granularity.** For some domains of reporting, there was broad agreement that collecting more granular data – underpinned by clear, standardised definitions – could reduce ambiguity of instructions and allow data to be aggregated for more than one report.

However, some larger firms expressed concerns relating to quality assurance and the technical challenges of supplying granular data, such as the cost of providing granular data from legacy systems and maintaining access to historical datasets. For example, the current



verification process may involve approving a small number of data points on a spreadsheet, whereas a future sign-off could involve quality assuring a data point model. Some firms suggested this would require an update to the Senior Managers and Certification Regimes



In supplying granular data, firms had similar concerns to proposals that meant full automation of the reporting process. They worried about losing ownership of their data and losing sight of how their data is used in generating reports. They felt that supervisors that did not understand the nuances in their 'granular' data may draw the wrong conclusions about the health of their firms. Some firms suggested that this misunderstanding may lead to more, not fewer, ad hoc requests for clarification. Therefore, firms would still require transparency and assurance about how the Bank is using the granular data they submit.





## Figure 7: Mapping governance and architecture and other solution areas to the data collection process

Solution areas will have differing benefits for each step in the data collection process (shown below). They will also bring new costs (not shown).

Solution Area:	Data collection process:						
	Analytical question and collection initiation	Rule-making governance, approval and publication	Impact analysis and interpretation	Solution development and implementation	Execution maintenance and submission governance	Assurance and use	Change and review
<b>Governance and architecture:</b>							
Moving from push to pull					■		
Central service provider			■	■	■		
Collection of granular data			■	■	■	■	
<b>Additional solution options:</b>							
Using alternative data sources				■	■	■	
Changing mix of regular and ad hoc collections					■		■
Better alignment with business processes and outcomes	■	■				■	■



### 4.4: Other solution areas

As well as raising additional challenges to some solution areas, feedback during the review also provided a number of new learnings. In particular, the review highlighted new solution areas we may need to explore as part of our transformation work. These included:

- **Using alternative data sources.** There was strong support amongst wholesale banking firms for the Bank to consider sourcing some of its data from industry intermediaries, such as financial market infrastructure, rather than directly from firms. This could result in fewer 'sources of truth' and a higher quality of data. However, from the Bank's perspective, this should not detract from the need for firms to 'know their data'.

- **Changing the mix of regular and ad-hoc data collections.** Some of the data the Bank collects regularly is primarily used for event-driven or ad-hoc purposes. These types of collections could be viewed as precautionary collections: data that is collected so the Bank can quickly gain policy or prudential insight when the need arises. Some members of the Bank's internal working group supported a shift towards a smaller set of data on a regular basis, supplemented with a better, more flexible, ad-hoc collection process to meet new reporting needs where necessary. Others took the opposite view, arguing for more robust streamlined regular collections that reduce the need for so many ad-hoc requests.
- **Better alignment with business processes and outcomes.** Why we collect data impacts 'what' data we collect and 'how' we design the data collection process. Many working group members felt this was crucial to help us estimate the value of a collection, and to help industry understand why we were collecting the data (see sub-section on Value and Collection Rationale). Internal working group members noted that the solution for data collection should align to the data's intended use. For example, if we do not use data to monitor firms in real time, we do not need to design a reporting process that collects real time data.



## 5: Establishing a vision for data collection transformation

---

The rest of this paper focuses on our vision for the future of data collection, and how we think we can work with industry to get there – our ‘transformation plan for data collection’. A series of industry working groups and events in the second half of 2020 has contributed to the development of this plan.

Our vision for data collection is that ‘The Bank gets the data it needs to fulfil its mission, at the lowest possible cost to industry’. Underlying that vision, our aspiration is that:

Data collection can help deliver richer insights into emerging risks, faster, and the Bank will be able to confidently and precisely act to mitigate those risks. In doing so it enables us to respond to unpredictable micro and macro-economic events (from individual firm issues to economic shocks like Covid).

- Data collection is lean, so we and industry can make the most of our scarce resource. A lean approach to data should also flag when data sets are no longer useful and should be retired.
- Data collection continues to be reliable, so data is collected and used securely, accurately and appropriately.
- The process and purpose of data collection is clear: clear to industry why we want data; and clear to Bank staff what we need it for.

We think a well-defined vision for the future of data collection is important in order to signal our direction of travel, and to start to sketch out the expected future state. Working group members agreed, with firms indicating that a clear high-level steer would help them commit resources and align relevant internal change work.

Working group members made several suggestions about what delivering the vision would feel like for firms. Their vision statements touched on a number of aspects of the process:

- **Merging external reporting with internal reporting.** Some members suggested that, in the future, their separate regulatory reporting roles and teams should no longer exist; instead, reporting should be fully consolidated within their everyday operating processes.
- **Agile reporting.** Some members saw the future of reporting as about the ability to rapidly deliver a reporting change. In this world the lines between setting up a new regular report and ad-hoc reporting become blurred.



- **A combination of solutions.** Some members felt a vision meant combining specific options discussed in our solution section. In particular some wholesale participants described a vision combining instructions as code, industry data standards, in addition to an improved industry and Bank engagement process.
- **A framework for open collaboration.** Some members emphasised that the vision meant deeper collaboration between us, firms, and their service providers. This deep collaboration was based on an eco-system of standards that allowed people and machines to talk more easily. Each standard could be developed independently to suit certain purposes or specific domains, but all standards should be developed and maintained under a common framework. The [European Commission](#) has explored this concept in the design of its ‘common data spaces’.

Members all agreed that, to achieve our vision, we need a fundamental transformation of how the data collection process works. To deliver this transformation, we think there are three crucial areas of reform: common data standards, modernised reporting instructions and an integrated approach to data collection.

## 5.1: Common data standards

We believe the bedrock of great data collection is industry data standards that can be used by all that need them; what we call ‘common data standards’. Common data standards are collectively adopted methods for describing equivalent operational data. These data standards should capture and describe heterogeneity in the financial sector – not try to hide it. These common data standards should be a first step in helping industry tackle the complexity they face within firms. In doing so, we think they can help improve all aspects of the data collection process: all the way from how firms first source data, to how we understand differences between data points that have been provided to us.

But we, and others we met, also recognise that common data standards are about more than just reporting to us. Standards are a key part of the soft infrastructure of the digital age. As such they can bring benefits throughout the financial sector: improvements in operational efficiency, greater clarity for firm management and firm investors, and support for new innovations like blockchain and artificial intelligence/machine learning methods.

Given those wider benefits, we see our push for common data standards for reporting as part of a wider move to standardise and digitise data and processes within the financial sector. We see that movement as starting with existing work on financial identifiers, including major wholesale market initiatives like the Common Domain Model project, as well as many other private initiatives in this space (see Box A).

But as with any infrastructure programme, some degree of coordination and leadership is required to make it a success. Without it, multiple conflicting standards may arise, standards may be developed but not adopted, or perhaps worse, standards may not be developed at



all. We intend to support the development of this infrastructure. Through our transformation programme and engagement with other initiatives, we aim to encourage and catalyse change (see chapter six). In doing so, we will act to ensure industry are not just developing standards, but the right standards; standards that will be fair to all users, and that will fit together to deliver the maximum benefit.

We recognise that it will not be possible to roll out common data standards in all areas. There may be instances where the complexity of firms' data may be too great, the costs of dealing with legacy too high, or the necessary international agreement unfeasible. In such cases, we commit to being pragmatic: for example, leaving operational data and systems untouched (and creating higher level common data inputs instead), only delivering change for new systems and products, or focusing on purely domestic solutions first.

## **5.2: Modernised reporting instructions**

We think a better data collection process also needs to reform how we write and publish reporting instructions, and how firms consume and interpret those reporting instructions. In doing so, we hope our instructions will be clearer and easier for their users to work with, and result in less variability in how they are interpreted. We see merit in continuing to explore versions of all the options for this that we discussed in the discussion paper and during our review process.

## **5.3: An integrated approach to data collection**

Common data standards and modernising reporting instructions are two reforms that can transform any given report or data collection. But to fully realise their value, and really tackle the complexity problems identified in chapter three, they should provide the basis for a more integrated data collection approach.

We see this integration reform having three dimensions. First, the data collection process should be consistent, regardless of which authority firms are reporting to, and what the data is being used for. For firms, this means breaking down barriers between external and internal reporting processes; for the Bank this means statistical, regulatory, and markets and banking reporting feeling like part of one consistent solution. Second, every step in the reporting process should be better integrated with every other step. So our use of data is taken into account when we write reporting instructions, which in turn take into account how firms interpret those instructions and supply the data, and finally how we store the data and make it available to its users. And third, that the data itself is integrated: both at the input level, within firms, and in the final data sets that we receive. So, for instance, for firm data, all residential mortgage data, regardless of its source system, should be presented in a consistent fashion, with similarities between residential mortgage data and commercial loan data also captured.



Integrating data should prevent unnecessary duplication of data, make it clear how different data points relate to one another, and reduce the complexity and cost of data collection for firms and the Bank alike.





## Box A: Learning from others

In order to develop these solutions, we are not attempting to start from scratch. There are successful initiatives from other industries, ongoing initiatives within the financial sector, as well as ambitious firm or vendor-specific projects to learn from.

**Other industries.** Many participants talked about their experience standardising operational data to improve reporting in other sectors. One participant talked about a process to standardise data from medical devices, and how that led to more consistent reporting and better industry analytics over time<sup>[19]</sup>. We think innovations in similar data-rich sectors such as pharmaceutical, food processing, telecommunications, aviation and law, can offer valuable examples.

**Derivatives industry.** The International Securities and Derivatives Association (ISDA) has orchestrated the development of the [Common Domain Model](#) (CDM) which offers a standardised approach to describing events and data for over-the-counter derivatives products. The CDM aims to be integrated into three key derivatives processes: the legal agreement between the parties; on-going operations such as margin calls and settlement; and regulatory reporting. The CDM is now being expanded to include securities lending products.

**International and historic regulatory initiatives.** As we laid out in the discussion paper, the Bank has been an active adopter of [XBRL](#) and Data Point Modelling techniques in its regular regulatory reporting. We have been involved in extensive international harmonisation efforts, including promoting key financial identifiers Legal Entity Identifier (LEI), Unique Product Identifier (UPI), Unique Transaction Identifier (UTI), and Critical Data Elements (CDE) for derivative reporting. We also have much to learn about how the [SWIFT](#) payments framework was created and rolled out internationally, and more recently, from the [ACORD](#) international messaging standards initiative for the insurance industry.

**Firm- and vendor-specific projects.** As we learnt during the review, there is a desire from industry to embrace new technologies to upgrade systems and improve reporting. For some larger firms, who carry out reporting on behalf of a large number of entities, these intra-organisation reform efforts look similar to the work we are proposing at an industry wide level. Specifically, firms discussed how they created and implemented a vision for data, their approaches to consolidating data estates, their experience from increasing automation within their firms, and their use of cloud computing to improve processing capability.



Technology solution providers are already starting to facilitate the use of common data inputs across multiple regulatory requirements, clearly seeing the benefits of a standardised process that enables reporting firms to map their internal data to external reporting data points.



## 6: Approach to data collection transformation

---



### 6.1: Delivery approach principles

We think the vision and associated reforms laid out in chapter five are valuable, feasible and necessary. That being said, we also recognise these reforms are ambitious and touch upon large areas of the UK, and global, financial sector.

We also recognise a large part of the cost of this change will fall on industry. The Bank is committed to providing resource for the design and delivery of these changes. But, as with current data collection changes, we expect that firms will need to invest multiples of what we invest in implementing these reforms<sup>[20]</sup>.



So, during the review process, we thought hard about how we tackle these reforms: how can we give confidence to budget holders that these changes will be worthwhile, how can we break down the reforms into manageable chunks, and how can we minimise the cost and disruption of change?

To respond to these challenges, we are adopting with a delivery approach based on four key principles. The change will need to be:

- **Long-term and vision-led**, so that those affected have time to plan, invest and sequence, where possible, with existing change programmes. We set out our vision and accompanying reforms in chapter five. We expect those reforms to require at least the ten years envisioned in the original review scope.
- **Incremental and pragmatic**, so that we can start small, learn, and overcome key challenges. As we learn we can then tackle larger-scale change, while at the same time quickly readjusting if things aren't working. An incremental approach was strongly supported by both internal and external working group members. Experienced change managers highlighted how similar complex projects had failed when they tried to do too much too quickly. In their view, disappointments are common with a 'big-bang approach' as these tend to be quite risky, take too long, and often end up with sub-optimal solutions that do not deliver true value.
- **Open and collaborative**, so that we can build on the work of others and they can build on ours. We don't expect to deliver all the reforms alone or within a single programme of work that we lead. Rather, by being open and clear about the problems we face, and our possible solutions to these issues, others can invest in solving those problems too, and we can reuse their solutions for the benefit of the sector overall.
- **Continuous value and use-case driven**, where 'use cases' allow us to manage scope – limiting each project to a defined report or set of related reports. Delivering value for each

use case builds confidence with the Bank and industry stakeholders that the transformation plan is working. As we prove value for one use case, this will support the investment case for future use cases.

## **6.2: Transformation programme**

To help deliver these reforms we, jointly with the FCA, want to set up a multi-year transformation programme. We want this programme to be delivered in line with the approach and value framework described above.

### **Transformation programme phases**

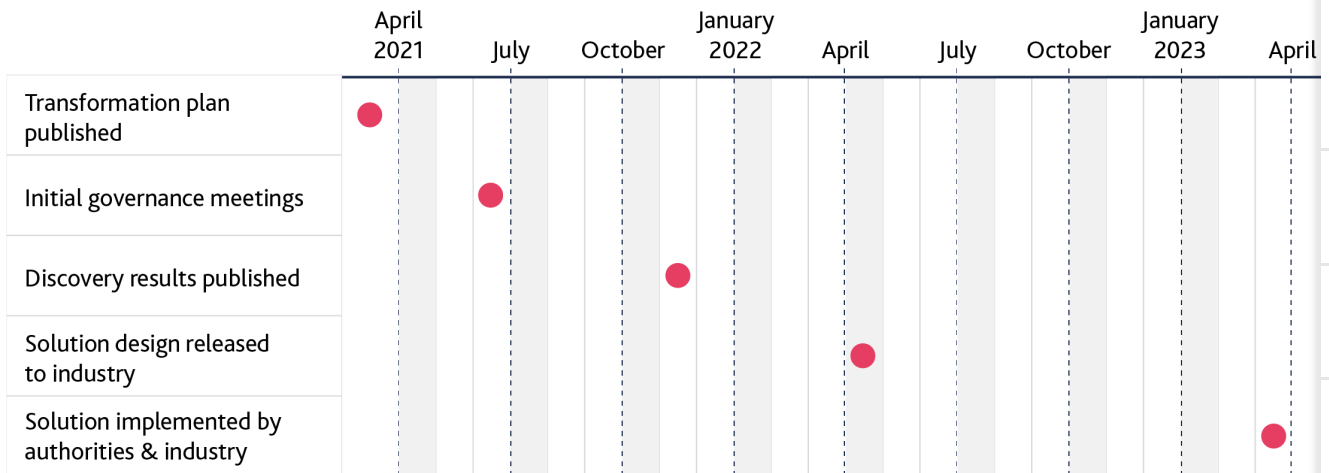
We expect our transformation programme to have at least three phases.

During phase one we want to set the foundation for future phases. Concretely, we think phase one will have three high level aims. Firstly to produce iterations of design prototypes, and to translate these into early versions of functional solutions; secondly, to deliver tangible change, for a limited number of use cases, thus creating value and showing stakeholders that transformation can be achieved; and third, to deliver intangible benefits, such as learning, key relationships, and establishing the teams and structures that will manage the programme. We expect the first phase to focus on making progress on the 'modernising reporting instructions' and 'common data standards' reforms, with progress on the integration of data collection to come in later phases.

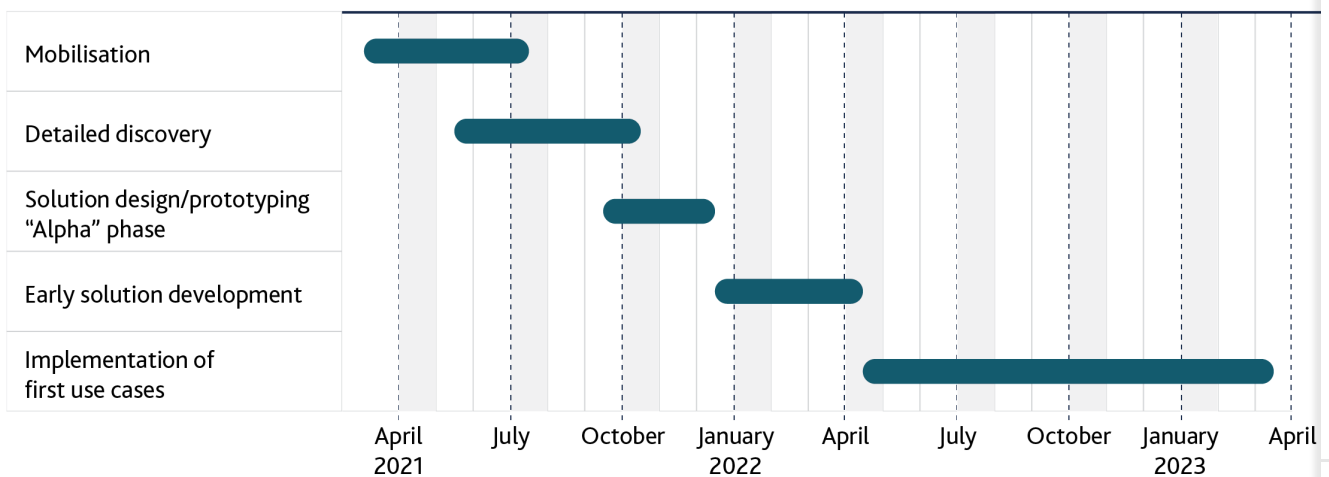


**Figure 8: Indicative phase one plan: 2021-2023[21]**

**Key milestones**



**Joint transformation programme**



During phase two, we will focus on scaling the transformation into new areas. Critically, we will look to scale the work we did in phase one and begin the process of integrating reports and data – key to our ‘integrated reporting’ reform.

During phase three, we will look to expand the work to more complex collections, building on the techniques we have developed and the results already delivered. We think these later phases will be where the bulk of the value will be delivered.

**Phase one delivery resources**

While we are aware that the global pandemic poses unique challenges to us all, we do want to start phase one this year. In particular, we are looking to build a core team comprising staff from the Bank, the FCA and staff from the firms from whom we collect data. We currently estimate the team’s size to be between 10-15 people, of which 12 might come from firms. We



expect the core team to be supplemented by part-time input from other staff from across the Bank, FCA and industry. We will need staff who have expertise and vision to contribute to designing solutions. In addition, to ensure the new changes work for everyone, we would like individuals who are directly involved in the current data collection process to provide feedback on issues and solutions.

## Phase one workstreams

During phase one, we expect the work of the programme to progress through two separate workstreams:

- **Common data standards.** Under this workstream, the programme team will look to identify and address issues relating to the consistent development and adoption of common data standards. As a first step, the Bank plans to set up an industry committee to work on UK financial sector data standards. We hope the committee will include a cross-section of industry representatives and relevant trade associations and standards bodies. We think its longer-term aims should include preventing fragmentation in the development of private standards, and acting as a single voice to the Bank and UK public authorities on actions to help standards adoption. We hope to have the committee established by mid-2021.
- **Reporting.** Under this workstream, the project team will aim to identify issues with reporting and design and deliver solutions. This workstream will include work on modernising reporting instructions. The initial activities of this workstream will be to create a common, detailed understanding of the problems faced by users of parts of the reporting process relevant to the use cases in question (see next subsection). This will set the foundations for solution design and the 'alpha' prototyping phase, before the agile delivery and implementation of solutions.



## Selected phase one use cases

We have identified three prospective use cases which, in our view, are representative of the Bank's data collection activities, and which we expect the transformation programme to look at during its first phase (see box B for an overview of the use case selection process). These use cases will sit alongside further FCA use cases:

- **Reform the quarterly derivative statistics return.** The Bank collects and [publishes summary statistical data on the UK derivatives market](#). Reforming this return allows us to align with, and build on, existing industry and global authorities' work on data standards for derivatives reporting.
- **Deliver a commercial real estate (CRE) database.** The CRE industry have a long running project looking to create a database on the CRE market<sup>[22]</sup>. We hope delivery of this project will help us meet our reporting needs. Aligning with this project will allow us to

develop common data standards in two core areas in the financial sector: loans and property. This in turn will help improve the quality of data in a market that is important to monitor for financial stability purposes.

- **Optimise the liquidity monitoring metrics (LMM) tool.** The Prudential Regulatory Authority's (PRA) LMM tool is an algorithm, published in Excel, that shows firms how the Bank calculates some key liquidity metrics using data defined by the [PRA's liquidity report PRA 110](#). Upgrading the tool provides an opportunity to understand how we might deliver 'instructions as code' – potentially a key component of a more flexible reporting process in future.

Alongside these core use cases we will carry out preparatory work on some 'high-value, high-complexity' use cases, where we believe we can integrate a series of reports. Mortgage data reporting is one such use case. We currently collect a number of mortgage reports, but the governance of mortgage reporting is complex and so initial progress will be slow. We expect to start delivery of changes for mortgage reporting and similar use cases in phase two of the transformation programme.

### 6.3: Alignment with other initiatives

Our approach to delivering reforms will not be developed in isolation. We expect the joint transformation programme to learn from, and align with, other relevant initiatives (see Box A for a sample of such initiatives). Specifically, from a Bank perspective, we intend to take a number of actions to progress the vision and reforms outlined in chapter five that will sit outside the joint transformation programme.

Internally, we will progress work to streamline internal processes and review what data we collect. In doing so, we will think about whether changes to our data collections would be good use cases for our transformation programme. For instance, for insurance reporting, the PRA is carrying out a full review of regulatory reporting, including Solvency II returns. The result of that review may lead to new use cases for phase one of the transformation programme. The Bank and FCA have on-going projects investing in the systems we use to receive and store data, and looking at improvements to our ability to collect data on an ad-hoc basis. We are currently integrating aspects of our regulatory and statistical reporting. This includes building a common model for statistical reports and implementing a machine readable taxonomy using the eXtensible Business Reporting Language (XBRL).

We recognise the importance of promoting harmonisation of data and reporting issues at an international level, and we continue to engage at multilateral fora and through bilateral channels on these issues. For instance, we contribute to international efforts at the Financial Stability Board (FSB), where we have been leading work, in collaboration with other FSB members, on potential ways to facilitate convergence in data reporting in order to address market fragmentation. We also continue to contribute to work that promotes the standardisation, development and adoption of key financial identifiers and derivative





standards. Moreover, we will continue to collaborate on relevant international innovation projects such as a joint project[23] run by the Singapore BIS-Innovation Hub to create common data standards and machine executable reporting rules for mortgage reporting.

We have been working with and under the governance of international standard initiatives to progress common best practice while collaborating globally. This has included our approach to the adoption of the ISO 20022 financial messaging standard for sterling high value payments.

### Figure 9: Our use case based approach to incremental delivery



#### We grouped use cases into roadmaps



##### Event driven and ad-hoc

- Crisis lending scheme data
- Solvent wind-downs
- Ad-hoc collections



##### Product based collections

- Consolidated loan collection
- Quarterly derivative statistical return
- Commercial real estate database



##### Prudential regulatory

- Liquidity monitoring
- Asset encumbrance

#### And answered these prioritisation questions



##### What might we expect to deliver

###### Value

What value might we expect this use case to deliver?

###### Transformation potential

Can this use case help the programme as a whole?



##### What do we need to deliver?

###### Feasibility

How easy is it to try new processes for this use case?

###### Resource requirements

How much resource do we need to deliver changes in the use case?



##### What are we willing to give to deliver it?

###### Risk appetite

What are the impacted areas' appetite for change?

###### Resource availability

How much resource can impacted areas commit?

### Box B: Use case overview

A critical component of the plan will be the 'use cases' that we look at in the first phase of delivery. These use cases will shape the scope of a more detailed problem definition and solution design process from 2021 onwards. Implementations of

solutions for these use cases will deliver the first items of value from the transformation programme, and set the foundations for the delivery of future value.

## Use case assessment framework

During the review we developed a framework for assessing use cases. The framework aimed to select and order potential use cases into a coherent road map. Our framework had three high level inputs:

- What value can we deliver: what value can we add, and how will the use case help our overall transformation goals? This included what value a use case might deliver to the stakeholders involved in that use case, and what value the use case might deliver to a transformation programme as a whole – what we called ‘transformation potential’. Transformation potential in turn had two perspectives, both of which could add value to future use cases: tangible foundations (like data standards), and intangible momentum (like stakeholder confidence).
- What do we need to deliver: how feasible is the use case and what resources do we need to deliver it? This included the ease with which existing data ingestion mechanisms can be changed, timing can be aligned with other regulatory changes that are planned or in progress, or the ability to handle a large volume of data.
- What are we willing to give to deliver: what resources are available and what is our risk appetite in instituting change for the relevant data collection? When discussing use cases, working group members were aware of the impact COVID-19 was already having on short-term resource availability in industry and within the Bank.

## Use case roadmaps

In line with our incremental approach, we expect the priority of use cases to change over the course of the transformation programme. We think some use cases will be better suited to experimentation, developing new solutions and require less resource. These will be better use cases to deliver in earlier phases – when we will look to develop new solutions and prove their value. Other use cases will be more complex, but higher value, and are better suited to be delivered in later phases. Finally, we believe that use cases will have commonalities that mean delivering one use case will provide the basis for the delivery of a future use case – they have high ‘transformation potential’. We think creating these ‘use case roadmaps’ can help mitigate the primary risk of a use case-driven approach: that the delivery of an individual use case does not aid the delivery of future use cases.

We developed three use case roadmaps where we think there may be high commonality between use cases:

- event-driven and ad hoc use cases (triggered by external factors or firm-specific issues)



- product-based use cases (collections covering data sourced from a specific business or product line)
- prudential / regulatory use cases (regular collections aligned to key regulatory frameworks such as banking capital or resolution regulation)

## Use case selection

During the review, participants made suggestions for over 50 possible use cases to consider. The use cases differed greatly in their complexity and ambition, reflecting the diversity and scope of the review. Some use cases focused on data submitted by retail and commercial firms, others on data submitted by wholesale. Use cases differed by primary purpose – with regular supervisory and statistical use cases captured alongside event driven reporting. In addition, use cases differed in terms of complexity and the ease with which a project can be framed and implemented in the near future. For instance, some use cases had significant dependencies and external factors that made short-term delivery unfeasible.

This initial list of use cases was whittled down to a longlist of sixteen use cases by the Bank. These use cases were assessed against the use case framework, discussed at working groups, and voted on by working group members. The feedback on certain use cases was often mixed, as working group members weighed trade-offs between value and feasibility. For instance, some participants considered the PRA's key liquidity report, PRA110, a good use case because it is based on multiple cuts of the same data, and is complex and costly to execute. Others felt its complexity made it unsuitable as a use case – at least in the short-term. A selection of these use cases is included in Annex 4.

The working group use case evaluation fed into our assessment of which use cases would be most suitable as a first implementation. When deciding on the final three use cases however, we also took into account our internal assessment of feasibility. This meant some use cases favoured by the working group members weren't possible in the short-term. Some of the use cases ranked high in value by working group participants required too much prior governance work to be delivered within a year. Others clashed with upcoming policy changes.



# Annexes

---



## Annex 1: Overview of the review process

### The Discovery Phase

- In addition to analysing written responses, we held a number of events, engaging with a range of direct stakeholders. This included 75 different bilateral meetings as well as roundtables, webinars and conferences. The majority of the discussions took place during bilateral conversations solely aimed at providing in-depth feedback to the discussion paper. At these events, we talked to a broad range of organisations – regulated firms, private authorities, solutions providers and public authorities – often with senior leaders in data and reporting (see figure 1).
- We also had internal discussions within the Bank, with colleagues invited to suggest potential data collections that could be improved. These yielded in excess of 50 ideas across regulatory, statistical, markets and other collections.

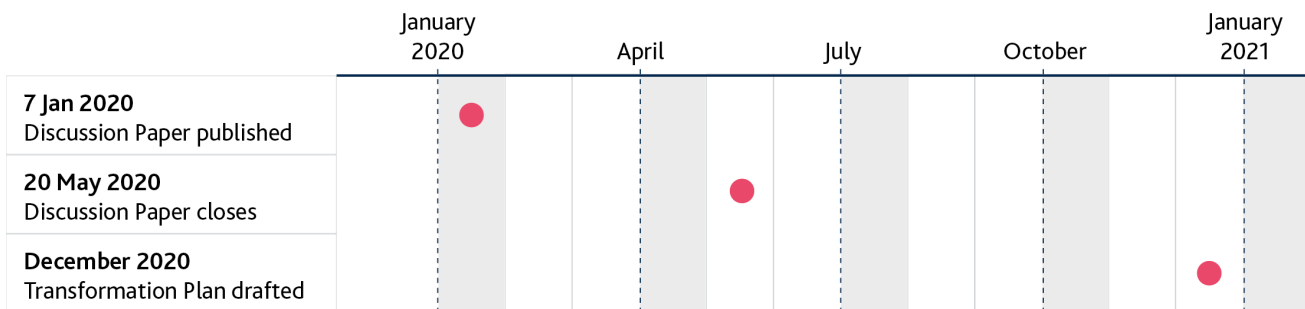


### The Shaping Phase

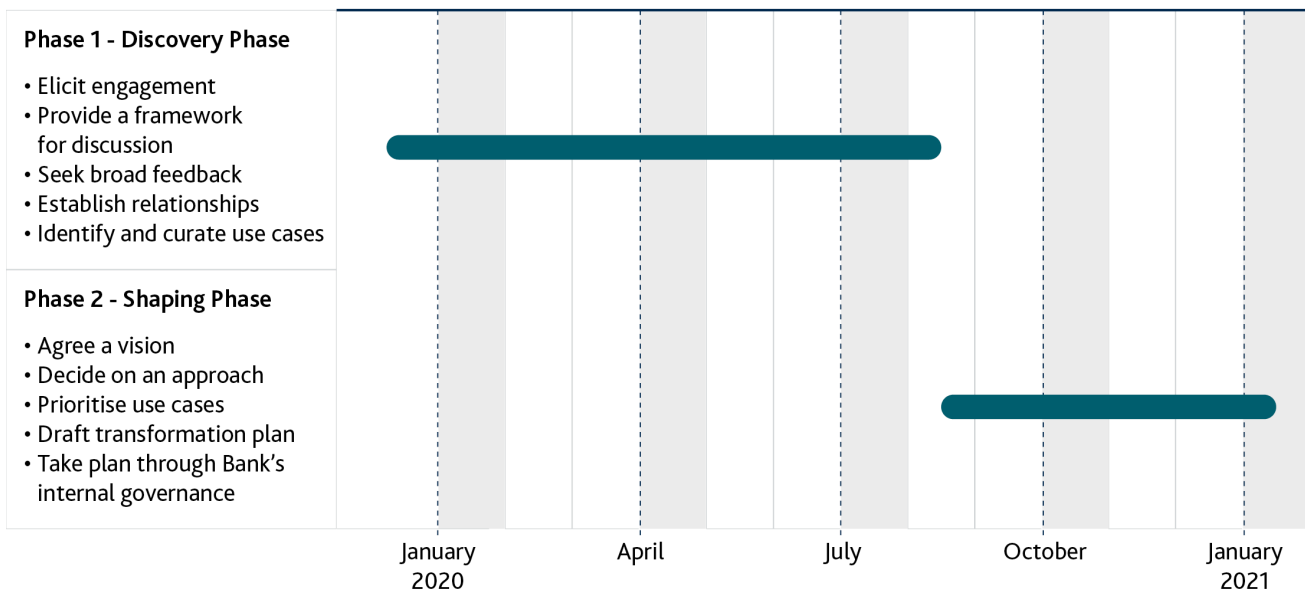
- From June to October 2020, our focus turned to shaping the details. Having made a call for nominations along with the launch of the discussion plan, we selected representatives from industry to participate in three sets of working groups: retail and commercial; wholesale; and insurance. The list of firms represented is listed in Annex 2. These were complemented by additional events such as vendor roundtables and bilaterals with a wider range of stakeholders (with a particular focus on regulators, data standards bodies, solutions providers and so on).
- The initial list of data collections that could be improved was whittled down to a selection discussed at the working groups.

## 2020 Review Timeline

### Key milestones



### Review phase



## Annex 2: List of external working group participants

### List of external working group participants

Association of British Credit Unions (ABCUL)

Association of British Insurers (ABI)

Association of Foreign Banks (AFB)

Association for Financial Markets in Europe (AFME)

Ageas Group

American International Group (AIG)

Aldermore Bank

Aviva

AXA XL

Barclays Bank

BNP Paribas

BNY Mellon

The Building Societies Association (BSA)

Citibank UK

Coventry Building Society

Credit Suisse Group

Cynergy Bank

Deutsche Bank

Direct Line

The Family Building Society

The Futures Industry Association (FIA)

Goldman Sachs

HSBC UK

International Capital Market Association (ICMA)

Investec

International Swaps and Derivatives Association (ISDA)

International Securities Lending Association (ISLA)

JP Morgan

Just Finance

LCH

Legal and General



Lloyds Bank  
Lloyds of London  
Macquarie  
Monzo  
Morgan Stanley  
NatWest  
Nationwide Building Society  
Nottingham Building Society  
Phoenix Group  
Pension Insurance Corporation  
Prudential  
Royal London  
RSA  
Santander  
Sumitomo Mitsui Banking Corporation (SMBC)  
Societe Generale  
Standard Chartered  
TSB  
UBS  
UK Finance  
United Trust Bank  
Virgin Money  
Yorkshire Building Society



### **Annex 3: A stylised view of the data collection process[24]**



## Stylised phases of data collection

Reporting firms submit data collections to the Bank both as a regulatory authority and as a statistical collection authority. The seven phases described below are our attempt to separate out the steps involved in the reporting process. Although there are dependencies involved, the process is not strictly linear, and there may be some iteration involved, or some overlaps between steps. For example, during the consultation process, firms do a rapid impact analysis and interpretation in order to give the Bank constructive feedback on proposed new regulations. However, firms only do a deep dive into the content once the regulations are published in their final form. In this instance, not only is the move from one step to the next blurred, it also includes an iteration.



### Analytical question and collection initiation

The reporting process ultimately starts with a policy question. This may be to better understand a new risk, to help design or monitor compliance with a new policy, or to evaluate the impact of a new policy. The question generates a request for data from a business area within the Bank or PRA for us to collect data.<sup>[25]</sup> During this phase, we decide which data we want, how these are defined, when we want them, and from whom we will collect them. We also consider the practical details about data collection – how the data are going to be collected, what legal powers we will use (if any), and the systems and resources required.

#### Figure A1: Key questions in the analytical question and collection initiation phase



##### Authority

- What question do we want to answer with the data?
- What data do we need?
- How do we wish to define the data?
- Why do we need the data and how will we use them?
- Where / who (which source) should we collect it from?
- When do we need it?
- What do we need to do to collect the data?

### Rule-making governance, approval and publication

Before a new permanent regulatory data collection can happen, the Bank needs to review and approve the collection. The approval process differs depending on the legal powers that the Bank is using to collect the data.<sup>[26]</sup> The Bank will publish draft proposals, after further internal consultation and sign-off from a senior committee or person. For regular regulatory






collections, the Bank undertakes a statutory Cost-Benefit Analysis (CBA) process, which tries to weigh the value of the collection versus the expected cost of delivering the collection. The CBA also helps us to build the business case for the collection.

A public consultation on draft proposals follows, which will likely include both the requirements and the instructions for preparation of the data. After consideration of feedback, the Bank can decide to collect less data (or even none), collect it with a reduced scope or frequency, or collect the data as proposed. Internal governance processes approve the revised version for final publication.

**Figure A2: Key questions in the rule-making governance, approval and publication phase**



---




- Do we need the data at all (does the cost of collecting the data outweigh the benefit)?
- Are we allowed to collect the data?

**Authority**

- Are our requirements clear?
- Are we meeting all legal requirements in our process for collecting the data?

---



- Do we currently collect the data requested?
- What is our best estimate of the additional cost to our reporting processes?

**Firm**

**Impact analysis and interpretation**

The impact analysis is a crucial part of any industry review of proposed changes. When requirements and instructions are published, firms try to understand what they need to do (if anything), when they need to do it, and the impact on their firm. The impact analysis may be part economic ('how much will this cost?') and part operational ('what people, processes and systems are likely to be affected?'). Once finalised, firms expend more effort determining what has changed, what is new, and how it applies to them.

## Figure A3: Key questions in the impact analysis and interpretation phase



### Firm

- Are we in scope of the data collection?
- Which data do we have to submit?
- How do we produce the data?
- What business areas are impacted?
- What systems and process are impacted?
- What is the likely direct cost of implementation?
- Are there likely to be any indirect costs of implementation? If so, what are they?

## Solution development and implementation

On the industry side, firms build solutions that carry out those instructions. To do so, some firms may need to contract out to third parties. Irrespective of their source, solutions are likely to involve a mix of technology, people and processes that lay out how people and technology interact. A solution may include some or all of the following steps:

- Setting up a project to fund, govern and manage the work.
- Documenting their interpretation of how to apply the instructions to their business.
- Identifying and sourcing data to meet the request, potentially across multiple operational, risk, finance and other management systems as well as purchased third party data sources.
- Establishing and testing solutions to integrate, cleanse, enrich and aggregate existing data and to create and submit the reports in the format required.
- Setting up or adjusting review processes to check and sign-off the data for submission.

Implementation also involves work for the Bank. We have to create the solution to receive, store and use the data we are collecting. This will differ depending on the type of collection. But for regular collections, it will typically involve making changes to the data collection portal that receives the data (such as BEEDs, or RegData<sup>[27]</sup>), as well as databases to store the data. Internal participants felt the introduction of XBRL (eXtensible Business Reporting Language) standard and DPM (Data Point Modelling) has improved the efficiency of this process and organisation of data in recent years.

## Figure A4: Key questions in the solution development and implementation phase

---



- What changes do we need to make to systems/processes?
- How are we going to store the data?

**Authority** • Does the system / process work as expected?

---



- How are we going to source the data?
- How do we turn the raw data into data we can submit?
- What solutions and processes do we need to put in place to do that?
- How do we test that these processes are operating correctly?
- How will we verify the quality of the data before submission?

**Firm**

## Execution, maintenance and submission governance

On an on-going basis, firms need to carry out the processes set up in the solution build phase. Moving from the phase of building a solution to running a solution may take time. For new reports, firms often go through a few execution cycles before their reporting systems are bedded down and working well. Some regular ongoing collections, such as FINREP and COREP, are expensive to run. This is primarily because they require manual processes that collate summary data from a myriad of sources.

Firms will also continue to maintain and make improvements to data collection solutions. They may choose to invest in a new system or process; they may fix issues with the solution; or they may test to ensure external changes don't negatively impact systems that are used for multiple purposes.

## Figure A5: Key questions in the execution, maintenance and submission governance phase

---



- Does the changed system / process work as expected?
- What additional changes to systems / processes need to be made?

**Firm**

## Assurance and use

Data received by the bank is quality assured and made available to the relevant user of that data<sup>[28]</sup>. Assurance includes both validation – ‘is the data definitely wrong?’, and plausibility checking – ‘is the data extremely likely to be wrong?’ If these checks highlight that the data are wrong or likely to be wrong then firms may have to provide explanations for values they have reported or correct errors and resubmit. The Bank intends to extend its formal review processes to canvas for feedback from internal users of data, to confirm that the initial question that sparked the data collection has been adequately answered.

### Figure A6: Key questions in the assurance and use phase



#### Authority

- Has the firm submitted data that breaks validation rules and is definitely wrong?
- Do plausibility tests show patterns that would suggest the data are likely to be wrong?
- Do we understand the data we are using?



#### Firm

- Do we understand variations in the data we have submitted?
- Are we confident that the data are correct?
- What caused the errors in the data?
- How do we rectify this for resubmission?
- What design improvements should we make?

## Change and review

Over time, as technology improves, regulations change, regulatory interventions evolve, market practice changes and new data sources emerge, the utility of some data collections will diminish, and/ or be able to be refined or consolidated. Sometimes, collections need to be discontinued.

The Bank periodically carries out a data stock take, but a more formal process for reviewing collections on an ongoing basis will involve expiry dates on collections. In addition, the Bank will need to ask firms for their views on whether long-dated data requests still map well to their data sources.

## Figure A7: Key questions in the change and review phase



### Authority

- Is the data collection still relevant?
- How do we use the data?
- How should we change our collections?
- What does this mean for industry?
- What does this mean for our systems?
- How can the change be implemented at minimum/ low cost?



### Firm

- What changes to systems/processes need to be made?
- How can we use technology to improve our data processes?
- Do the benefits of technology investment justify the costs?

## Annex 4: Other possible early use cases

Several other early use cases were shortlisted by at least one working group. These are:

**Loan book data submitted by non-systemic supervised UK-focused firms.** The PRA collects a dataset on exposures to different kinds of loans – mortgages, asset finance, motor finance, invoice finance, credit cards, personal loans, overdrafts and other lending from non-systemic firms. The data is submitted via Excel spreadsheets which have limited validation and a complex structure. Re-designing these collections so they are more clearly defined and could be validated on submission will improve data quality and reduce the manual processes and costs for both the PRA and firms.

**Crisis lending scheme data.** Firms and the Bank can face large operational costs when setting up new collections during a crisis. Establishing a flexible process to quickly set up a data collection for future lending schemes (similar to BBLs<sup>[29]</sup> or TFSME<sup>[30]</sup> eligible loans) will enable the Bank (and industry) to respond faster to the next crisis without any degradation in data quality.

**Asset encumbrance.** The Bank collects a detailed breakdown of unencumbered assets by type and by quality as part of regular COREP and PRA 110 submissions. It can be difficult to reconcile the total amounts across different sections of the returns. Automated processes could flag inconsistencies and reduce the time spent correcting errors.

**Definitions in COREP<sup>[31]</sup> and FINREP<sup>[32]</sup> reporting.** The PRA collects similar data points on financial reporting submitted as part of COREP and FINREP reporting. These data points, depend on how individual firms interpret the terms in the rules and instructions. This affects

the comparability of data points. Applying a consistent definition to financial reporting concepts across COREP and FINREP reports will align data points to widely accepted international standards and make it easier to compare firms.

**Collection of net spread earnings.** Sterling and foreign currency 'net spread earnings' derived from trading in foreign exchange, securities and derivatives are a key measure of firm income used in GDP calculations and are collected as part of statistical reporting. However, firms measure these for their own accounts. Creating more consistent and accurate estimates based on firms' operational income will improve the accuracy and consistency of the 'net spread earnings' measure in GDP.

**Non-economic trade data in post-trade processing.** Non-economic trading data (e.g. identifiers such as Legal Entity Identifier (LEI), Standard Settlement Instruction (SSI), Unique Transaction Identifier (UTI), and trade allocations, product identifiers such as tickers/International Securities Identification Number (ISINs)) used by market participants in post-trade processing can suffer from data quality issues. These issues generate a number of pinch points throughout the trade life cycle where firms have to check data quality, enrich trade information, reconcile their own records with the records of their counterparty and deal with breaks. Improved standardisation, accuracy, and timely exchange of non-economic trade data between market participants could reduce the need for subsequent correction and enrichment of trade information. Not only could this benefit market participants directly, but it could also be of benefit to authorities who collect and use post-trade data to support their monetary policy, financial stability, and/or supervisory objectives.

**Mortgage data reporting.** Data on mortgages is received through multiple mechanisms, reported by varying reporting populations and collected for a variety of purposes by the Bank and FCA (for instance FCA submissions used by the Bank include Product Sales Data (PSD) and Mortgage Lenders and Administrators Return (MLAR) data, while Buy-to-Let and stress testing data are submitted directly to the Bank). Data can be of variable quality; reconciling data from different sources is difficult and plausibility checks involve a lot of staff time. Firms need to report the same – or very similar – data more than once. Creating a comprehensive high quality mortgage data that can be easily used for a variety of research, analytical and risk assessment purposes may decrease the time spent on plausibility and error checking. In addition it may improve the Bank's ability to reconcile stock and flow data sets across time series.

---

1. Carried out with the FCA and seven financial firms.

2. Phase one use cases are 'quarterly derivative statistics return', 'commercial real estate (CRE) database' and 'liquidity monitoring metrics (LMM) tool'. See chapter six for more information.

3. With the FCA, Monetary Authority of Singapore and the International Swaps and Derivatives Association.
4. For the purposes of this paper, this includes the Prudential Regulatory Authority
5. Note the grids represent the ratios of different types of organisations and not numbers. Annex 2 has a full list of organisations included in the working groups.
6. One-off collections have a similar progression, but some steps such as public consultation do not occur. Overall, phases are truncated and the process is much faster to design and implement.
7. We use the term authority to generalise the party collecting the data. We think the process is broadly consistent across regulators, central banks and statistical compilers – the three key roles we play in the data collection process.
8. In turn, these 40 consolidated liquidity reports are built from 500 entity level reports.
9. See paragraph 2.19 of the [Discussion Paper](#).
10. We undertake some collections on behalf of third parties to help them meet their purposes.
11. The Common Equity Tier 1 Ratio is a key regulatory metric of the health of a bank that aims to capture the possibility of losses a bank may sustain relative to its ability to safely absorb those losses.
12. See chapter seven of the [Discussion Paper](#).
13. This issue is explored in more depth in chapter seven of the discussion paper.
14. Internal business line, policy, finance and/or compliance functions.
15. Firm participants said often they use forums provided by trade bodies or solution providers as informal forums to come up with common interpretations of reporting instructions.
16. Discussion paper chapter four, paragraph 4.9.
17. A Skilled Person Review is one of the regulatory tools we can employ under FSMA 2000, to commission an expert ('skilled person') to conduct a review of a reporting firm, on subject matters of interest/concern to the [PRA](#).
18. Firms noted two reasons why this might occur. Firstly, that how financial products are implemented in systems can change over time. These differences in implementation can result in small differences to the payments associated with those products. Secondly, that new systems may not support legacy products. In both cases, building and maintaining a system that meets the need of legacy and current requirements is prohibitively expensive.
19. See [Federal Drug Administration's industry-wide initiative](#).
20. See the [Discussion Paper](#) for a further discussion of Bank and industry costs.
21. Note timeline may vary due to requirements of specific use cases and in order to help alignment with other FCA, Bank and industry initiatives.
22. The database was originally proposed by the [Vision for Real Estate Group](#) in 2014 (an industry group that came out of the Bank's Commercial Property Forum). The Bank expressed its support for the initiative, notably through a [speech by Alex Brazier](#) in 2015.
23. With the FCA, Monetary Authority of Singapore and the International Swaps and Derivatives Association.
24. This has been updated from a previous version presented in the DP, pages 16-18. In practice, the process will vary depending on the significance of the proposed collection.
25. See the discussion paper chapter two for a further discussion of what data we collect and why we collect.
26. See discussion paper chapter two for more information on the legal powers used to collect the data



27. BEEDS is the Bank of England's data submission platform, and RegData is FCA's equivalent that replaces the current GABRIEL system. RegData / GABRIEL are used to collect some regulatory banking returns on behalf of the PRA.

28. See chapter two of the discussion paper for a comprehensive description of why the Bank collects data.

29. The UK government's coronavirus [Bounce Bank Loan Scheme](#)

30. The Bank of England's [Term Funding Scheme for SMEs](#).

31. The commonly used abbreviation for 'Common Reporting', the European Banking Authority's capital reporting framework.

32. The commonly used abbreviation for 'Financial Reporting', the European Banking Authority's financial reporting framework.

©2023 Bank of England

---

