



---

# The FCA Data Strategy - How we will manage and use the data we collect

September 2013





# Contents

<b>1.</b>	Overview	2
<b>2.</b>	Our vision for the future of FCA data	4
<b>3.</b>	Our strategy to deliver FCA data	6
<b>4.</b>	How we plan to implement our strategy	12

# 1. Overview

Our data strategy sets out why data is important to us as a new regulator, how we want to collect and use data in the future, our approach to delivering our vision for data and, recognising the size of the task ahead, our plan for implementation. We have decided to publish this data strategy so we can share our vision and the challenges we foresee, as well as gather feedback to inform our approach.

## Why do we collect data?

---

As a regulator we are naturally interested in how the firms and markets we regulate behave. We engage in all aspects of financial services and frequently probe to further our understanding. This results in a wide range of questions, from the number and type of firms operating in a particular market, to the more complex issues of whether a market is functioning well. The starting point to answering our questions is often to examine the data and information that we hold. Along with other sources of information, we ask our regulated firms to report a wide range of data, which we rely on to fulfil our objectives as an organisation.

## Our data legacy

---

We recognise that in the past the Financial Services Authority (FSA) did not always request data and information from firms in a clear and effective way. We have listened to the firms we regulate, who told us there are a number of failings, including:

- Too many requests for data and information without a clear explanation about why it was needed.
- Unreasonable timescales, resulting in firms needing to divert resources to meet our requests.
- Failure to communicate what the data and information was used for, leaving firms questioning if it was used at all.

In addition to the concerns raised by firms, we recognise that the FSA:

- Did not have appropriate internal governance to identify all the requests being made of firms to ensure the appropriateness of requests, consistency of communication and the re-use of data and information already gathered.
- Did not have all the necessary technology to fully exploit data and information throughout the lifecycle of collection, storage and use.

This is the data and information legacy that we have inherited and we are committed to addressing it. However, overcoming this legacy will take time, as much of the data and information we hold and continue to collect was initiated by and for the FSA, and our current data handling capability was defined by it.

If you want to comment on or give us **feedback** on our strategy, please contact the FCA's Data and Analysis Department:

**Email: [fcadataandanalysis@fca.org.uk](mailto:fcadataandanalysis@fca.org.uk)**

We are also keen to hear from interested parties willing to help us develop and deliver our data strategy and hear feedback from firms about the data and information we are requesting.

## 2. Our vision for the future of FCA data

Data and information is a key enabler to our success, supporting our drive to be a forward-looking risk-based regulator that operates efficiently. Good quality data and information, handled well and available quickly will give us deeper insight into the markets we regulate and allow us to be more efficient at identifying and tackling risks.

Our vision for the data that we collect and hold is to ensure that it is **actionable, integrated and fully accessible across the FCA to help us meet our objectives.**

We need data and information to be at the heart of our operations as a regulator. In future our data will be:

- **Effectively governed and controlled**

We will have effective governance and controls from the point where we specify our data needs through collection and use to the point of disposal. We will make all users of data accountable with senior management ownership of specific data sets.

We will establish governance and controls to ensure that we continuously review the appropriateness of our data collections and stop data collection where it no longer meets our needs.

We will ensure that our communications, policy and rule-making processes for collecting data are sufficiently clear, flexible and adaptable to meet our needs while balancing the need for certainty from the suppliers of data.

- **Clearly specified**

We will ensure that we specify the data we need in the context of our objectives and strategic priorities and make clear to suppliers how we intend to use the data. Our governance and controls will challenge any data collections that cannot show a clear link to our objectives and strategic priorities or to our obligations as an international regulator.

Data requests will be clear and unambiguous in what is required of the provider and the format in which it should be supplied. Wherever possible we will use standard formats and technology to enable data collection.

- **Fit for purpose**

All regularly-reported baseline data will be sufficient to meet a broad set of regulatory needs and support the use of judgement by staff. In line with our desire to step in earlier and act faster when we identify problems, we will specify regular data, designed to give us a broad overview of the markets we regulate, collecting enough to enable us to act or seek supplementary data where we identify a problem.

- **Strong and rich at a baseline level**

Our baseline data (that is, the data on which we base our views of how the markets we regulate are operating) will be strong, predictable and sufficiently rich to enable us to oversee the firms and markets that we regulate on a forward-looking basis.

Our baseline will be sufficiently rich so we can identify trends and risks in markets that may have an impact on consumers at an appropriate level of granularity that we can form a view and take steps to remedy. Our baseline data will consist of data collected from a wide range of the most appropriate sources, including the firms that we regulate, for us to understand consumers, the products they need and buy and the actions of firms operating in markets.

- **Collected appropriately**

All data will be collected through controlled channels to ensure that it can be consolidated, catalogued, directed and accessed across the FCA. We will ensure that we speak to suppliers of data before we make a formal request. Collection mechanisms will be easier to use and flexible to evolve, without significant lead time and costs for the industry.

- **Managed and stored in appropriate technology solutions**

We will use fit-for-purpose technology to use the data we obtain to its fullest extent in the most appropriate manner. We will ensure that our technology investments provide value for money and are sufficiently flexible and adaptable to meet our changing needs. We will seek, wherever possible, to use up-to-date, off-the-shelf technology.

### 3.

## Our strategy to deliver FCA data – how we will get there

The task of tackling the FSA data and information legacy is substantial and must be approached in a controlled way that **limits the impact** on those who supply data and information to the FCA.

Given the **scale of change**, we may need to make significant investments in technology. We are therefore taking a cautious approach by delivering this in an **iterative way**, so that we can test and prove the changes we need to make before we commit significant resources.

The initial focus for our efforts is on **collecting data from firms and specifically regular data** (that is data collected more than once and/or from more than one source), as this captures a **significant proportion of the data requests** we make of firms. Although our delivery focus at this initial stage is limited to regular data it is intended that the framework set out in this document will be **applied to all data and information in time**.

We have recognised that the most significant changes we need to make are in our **culture and behaviours towards collecting and using data**. We must change the way we think about data and recognise it for the valuable asset it is in delivering against our objectives. This will take time and we have shaped our delivery strategy to enable this change. Delivery of the FCA data vision comprises three related steps:

1. **Governance and ownership** – establish effective controls, oversight and ownership of data across the FCA
2. **Better engagement with firms** – improve our interactions with firms on data
3. **Better handling of data for the FCA** – launch and roll out new capabilities to bring about the FCA data vision



What we mean by these steps is explained below.

### **Governance and ownership – establish effective controls, oversight and ownership of data across the FCA**

---

To deliver effective governance and controls over data, we have identified that we need to centrally coordinate all of our data-related activities within a user area of the FCA. In response we have recently established a dedicated Data and Analysis Department within our Policy Risk and Research division.

The Data and Analysis Department will be responsible for:

- managing our internal data governance policies and controls
- overseeing our data sharing and coordination responsibilities with the PRA
- challenging proposed regular data requests of firms and data we are already collecting
- maintaining quality and integrity frameworks for our data
- supporting and shaping user requirements for technology to meet our data needs
- developing a focal point for all FCA data collections
- providing a Centre of Excellence for Analysis across the FCA

This new department will provide the essential focal point and momentum to achieve our vision for FCA data. Working collaboratively across the FCA, the department will help to embed the cultural and behavioural changes needed.

While identifying the data we need will still be generated from across the FCA, the Data and Analysis Department will ensure there is appropriate challenge of requests and that they are appropriately governed. In time, as our new data management capabilities are put in place the department will be responsible for ensuring our data is collected and controlled in line with our strategy.

As a first step to making change happen, the department has published a set of guiding principles for FCA data which are now starting to be applied.

### Guiding principles

- Data is always traceable to a defined FCA need for it in line with our objectives
- Data collection is centrally coordinated within the FCA
- All data collected is stored in a searchable accessible store
- Data can be readily and flexibly accessed by users from the store
- Data, in any form, is a FCA asset
- All data has clearly defined owners in the FCA
- Data can be re-used for a variety of purposes
- Data is handled in accordance with the FCA Data Management Model
- Data needs to evolve to support our desire to be a forward-looking, intelligence-led, risk-based regulator

### Better engagement with firms – improve our interactions with firms on data

---

One of the changes we have begun to make is to ensure that we fully articulate our need for data in the context of our objectives and strategic priorities. We want to ensure that for all data we collect, there is a clear line of sight from our objectives and organisational priorities through to its purpose, and the underlying enablers and controls that govern its use. We have established an Information Governance Board, which oversees and approves all new requests for regular data.

Firms should begin to see from now on:

- better engagement before any request being formally made on the nature of the request, the data that we require and its purpose
- more clarity on the data required

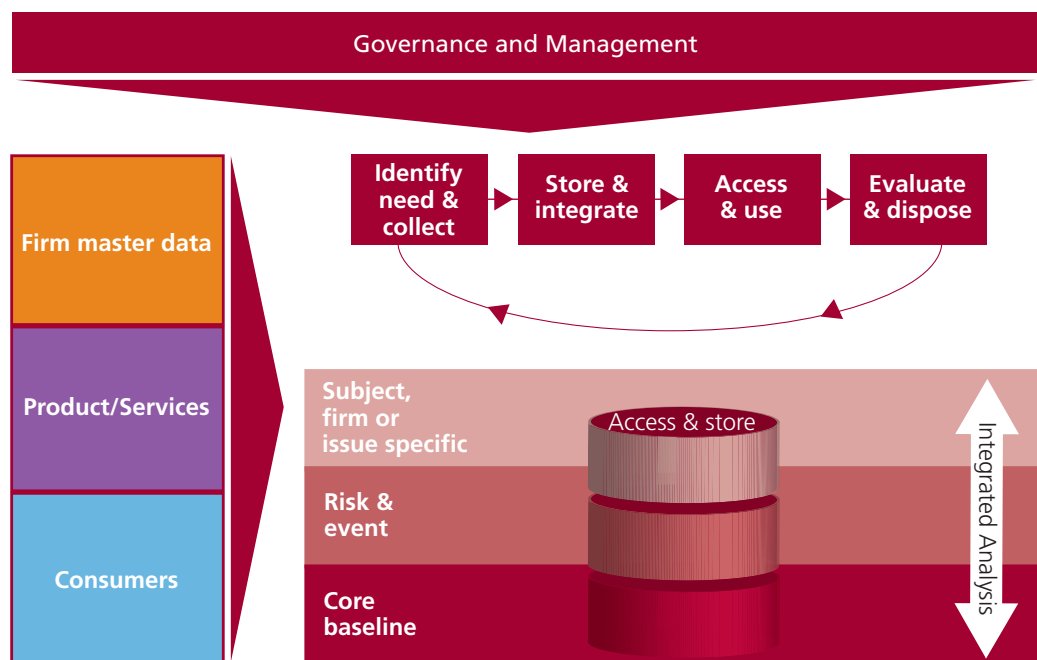
We are just beginning to make the changes we need to make with data; we would welcome the opportunity to work with firms and other interested parties to shape these changes. We are, committed to improving our interaction with firms on data as a priority and want to see significant changes in the course of the next year on how we articulate what we need, the volume of requests and their complexity.

## Better handling of data for the FCA – launch and roll out new capabilities to bring FCA data

To achieve the FCA vision for data we need to create a solid foundation, which gives users the ability to interrogate data regardless of its original purpose, mechanism of collection or storage. To do this we need our data to function in a joined-up way.

The diagram below illustrates how we need FCA data to work under our overarching governance frameworks described above. It comprises five key components:

1. The FCA Universe
2. Identify need and collect
3. Store and integrate
4. Access and use
5. Evaluate and dispose



## The FCA Universe

---

This represents the **single source of master data** about what we regulate and contains all the **static reference material** we hold about firms (such as **address, regulatory permission and contact information**), **products** (such as product design and distribution and our taxonomy against which we collect product data) and **consumers** (such as our consumer segmentation model).

This single source of master data will be the core component of drawing together FCA data in an actionable, integrated and accessible way.

## Identify need and collect

---

We need to collect data from a **wide range of sources** to meet our objectives. However, we need a simple way to **identify the data we need and why we need it** to ensure we collect it in a **proportionate, clear and fair way**.

We have identified three types of data that we need. Determining this from the outset will ensure we develop the right collection methods.

The first type of data is called **core baseline**. This data contains all the **regularly reported data** on which we base our regulatory activity. It is the core data that must be reported **regardless of any other data requirement** and should form the foundation for activity conducted by the FCA. Examples of core baseline data include firm returns submitted through the GABRIEL reporting system, product sales information and complaints information.

The second type is called **risk and event** and **supplements the core baseline**. Where risks are identified that form one of our strategic priorities, or where events occur that require data in addition to that held in core baseline, this is how we would collect such data. It is envisaged that data in the **risk and event layer** would typically **only be required for a period of time** while the risk is mitigated or the event managed.

The third type is called **subject, firm or issue specific**. This is one off data specific to a subject, firm or issue that we are acting on, but may be re-used in future (for example data collected as part of an **Enforcement investigation or through our authorisations activity**).

To implement the three data collection layers and log and direct the data into a **centralised store** that can be accessed across the FCA, we will need to **restrict collection methods** to a small number of **controlled compatible methods**.

### Store and integrate

---

Users of data within the FCA are concerned about having a simple way of accessing and linking all the data we have to provide an integrated view of a problem, risk or issue. Suppliers of data to the FCA do not want to have to supply the same information to a number of different parts of the organisation.

We will store the data we collect in an integrated way by cataloguing it when we collect it and ensuring our underlying data models enable us to identify corresponding data.

We will use solutions that enable us to store and integrate multiple sources of data without end users needing to understand the underlying mechanics of joining it up.

### Access and use

---

To enable data to be reused for a variety of purposes we will **make the data universally accessible** to all staff in the FCA. Data will be available for sharing through a **single point of access** for all users **unless there is specific market sensitivity**.

### Evaluate and dispose

---

We want to ensure that the data we collect is designed to meet our needs as a regulator, therefore we intend to **continuously monitor and evaluate its usefulness** and, where it is no longer useful, stop collecting it. We will then **remove and archive data** where there is no ongoing need to retain it.

These steps require a significant **amount of activity to deliver**. We do not envisage trying to tackle and deliver all of these changes at once. Our implementation plan is set out in Section 4.

## 4. How we plan to implement our strategy

Given that the size of the task to deliver this data strategy is significant, we are not planning to tackle all of our challenges at one time. Instead we are planning to deliver these changes iteratively and transparently.

We have three immediate priorities:

- **Governance and ownership** – establish effective controls, oversight and ownership of data across the FCA.
- **Better engagement with firms** – improve our interactions with firms on data.
- **Better handling of data for the FCA** – launch and roll out new capabilities to bring about the FCA data vision.

Our Data and Analysis Department, along with the Information Governance Board, intends to embed effective controls, oversight and ownership of data over the course of the next year.

We have a plan in place to ensure that all requests for regular data are approved by the Information Governance Board by the end of 2013 and we will be developing controls and measures to ensure this is the case. As part of this implementation, we will be working with the teams across the FCA issuing regular data requests to ensure they have a clear dialogue with firms about the purpose, format and timing of data requested.

### **Proof of concept of the new capabilities to bring FCA data**

---

A proof of concept is the demonstration of an idea to test whether it is feasible and beneficial.

We have a plan in place to deliver a proof of concept that will demonstrate all aspects of the data management model using a sample market before we make significant investments in technology. We intend to deliver, by September 2014, a working example of FCA data in the new data management model, which we can use to test our strategy.

We have chosen the UK Retail Investment Market because of the wide range of consumers. It is also evolving given the implementation of RDR and is sufficiently complex to test our integration capabilities.

Overall the proof of concept will test that:

1. we can **change the way we think** about data for the FCA
2. we can **identify and collect some baseline data** for the FCA to better understand how the UK Retail Investment Market is operating
3. we can deliver some of the enabling tools to access and use the data
4. we can design a repeatable approach for other sources of data and other markets

### Next steps

---

This strategy clearly sets out our vision and our approach towards addressing the data legacy that we inherited from the FSA and improving the way we manage data across the organisation. However, it will take time to achieve what we want to. We will continue to keep you updated as our thinking progresses and our plans develop.

If you want to comment on our strategy or give us **feedback**, please contact our Data and Analysis Department:

**Email: [fcadataandanalysis@fca.org.uk](mailto:fcadataandanalysis@fca.org.uk)**

We are also keen to hear from interested parties willing to help us develop and deliver our data strategy and hear feedback from firms about the data and information we are requesting.

**Financial Conduct Authority**



© Financial Conduct Authority 2013  
25 The North Colonnade Canary Wharf  
London E14 5HS  
Telephone: +44 (0)20 7066 1000  
Website: [www.fca.org.uk](http://www.fca.org.uk)  
All rights reserved