IFC-Bank Indonesia Satellite Seminar on *"Big Data"* at the ISI Regional Statistics Conference 2017

Bali, Indonesia, 21 March 2017

Data as a critical factor for central banks[1]

Maciej Piechocki and Anne Leslie-Bini,
BearingPoint / Central Banking

---

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

Data as a critical factor for central banks

Dr. Maciej Piechocki, Anne Leslie-Bini, BearingPoint

„ … it should also be clear to everyone that we are now standing only at the start of a long road in terms of data. The big challenge for statistics in the coming years is not only "many more numbers", but perhaps much more so, the reconciliation of statistical information collected in support of monetary policy and financial stability with the up-to-now rather separate world of supervisory information. It is one thing to have information, which, like blood, flows through the veins of the system, it is another to ensure that everything beats at the same rhythm and all organs in the body get all they need from the same single flow."
(Mario Draghi President of the ECB, Seventh ECB Statistics Conference "Towards the banking Union. Opportunities and challenges for statistics", Frankfurt am Main, 15 October 2014")

Abstract of conclusions and survey results from the Central Banking Big Data Focus Report

Central banking statistical stability or supervisory function have been increasingly driven by (big) data, but little has changed in the methodology of supervisory data collection and management, which is still widely reliant on the document-oriented approach. This is intrinsically time-consuming, costly and complex. Data gaps still exist and so data remains a critical factor for central banks. Innovative solutions are necessary, to effectively handle "Big Data".

The Central Banking Big Data Focus Report is a joint initiative of the Central Banking Journal and BearingPoint. The report builds upon the results of the recent IFC survey and takes a closer look at how central banks actually handle the challenge of data collection and analytics with regard to technical platforms and standards, resources and data governance.

The report investigates the concrete action plans of central banks regarding data management challenges in light of FinTech/RegTech developments and the objective of transparent and effective risk-based supervision but also plans for central banks statistics for "going beyond the aggregates" especially for the micro-granular data handling. Finally, central banks are evaluated how the BCBS 239 principle in an adapted version would apply to them today.

The focus report will draw on views from central bankers, industry experts, academics and observers to look at:

- Financial stability and supervisory applications
- Direct uses in economics and modelling
- Who should 'own' big data?
- Resourcing and budgets
- Future developments
- Operational challenges – gathering, structuring, storing and processing data

The Central Banking Big Data Focus Report aims at giving a clear picture of where central banks stand today with supervisory data management and defining fields of action.

Our part in the report sets out the results of a survey of how central banks view big data and data governance in their institutions. The survey was conducted by Central Banking Publications, in association with BearingPoint, in August and September 2016. The work has only been possible with the support and cooperation of the central bankers who agreed to take part. They did so on the condition that neither their names nor those of their central banks would be mentioned in the report.
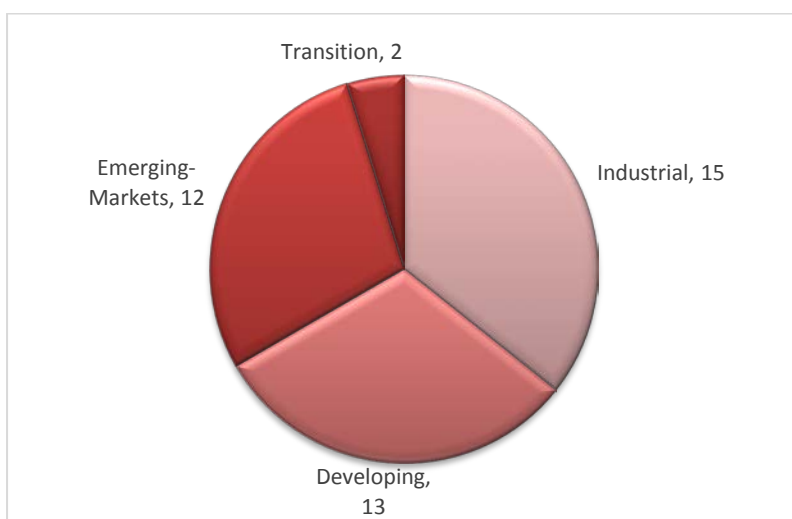
Summary of key findings[1]

- Central banks have an active interest in big data. This is manifested in improving processing technology, adapting institutional strategies and increasing staff awareness of the area.

- Central banks typically see big data as unstructured data that is sourced externally, though this view is not universally held.

- Overwhelmingly, central banks develop their own data platforms to handle regulatory data collection, a role that has taken on greater significance since the financial crisis as central banks have expanded their involvement in financial stability.

- Big data is predominantly regarded as useful for research, but significant minorities see immediate involvement in policy-making, or scope for this.

- Lack of support from policy-makers is seen as the most significant challenge to increase use of big data.

- Central banks do not in the main have a dedicated budget for the handling of data (including big data), though many are seeking one.

- A little over 80% of respondents said they do not have any intra-departmental or divisional bodies dedicated to big data.

- More broadly, central bankers have concerns over the arrangements in place for managing data in their institutions. Many are looking to improve data governance.

- Over three-quarters of respondents indicated they had a shared internal platform, typically in the form of reporting frameworks and data warehouses.

- Central banks generally source their own big data sets, though a significant minority increasingly look elsewhere for these. Overwhelmingly, they process these themselves and there is no indication of a desire for this to change.

- Monetary policy is seen as standing to benefit most from big data, though it is expected to have a significant impact on macro-prudential policy as well.

- Support from the executive-level and policy-makers divided respondents: 35% saw it as top priority for investment within the central bank, 38% saw it as the lowest.

- Central banks have developed their own data platforms to deal with regulatory data analytics, often used in conjunction with other options. Excel remains popular.

- Central bankers broadly welcome the idea of self-assessment of data management using an adapted version of the Basel Committee's BCBS 239 principles for supervisory data aggregation.

[1] Central Banking Publications, Big Data in Central Banks: 2016 survey results by Emma Glass

Article[2] content

Responses to the Central Banking Big Data Focus Report were received from 42 central banks al across the globe. The average staff size was 1,652 and three-quarters of respondents had less than 2,000 employees. Just over half of respondents were from central banks in Europe. Those individuals taking part in the survey were drawn in the main from the statistics function: 32, or three-quarters of respondents, were located in this area. Three were from information technology, and responses were also received from research, banking supervision, international relations, data management, infrastructure and technology, and data collection departments.
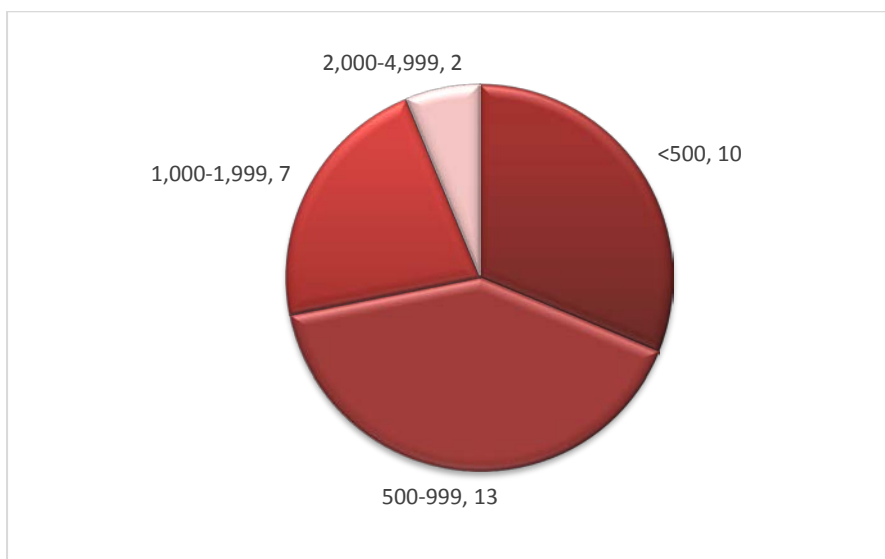
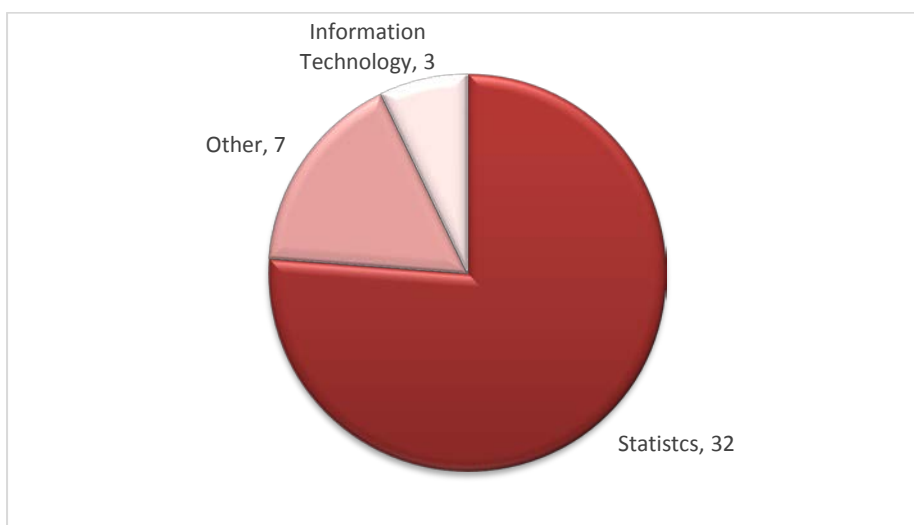Economic classification

Geography

---
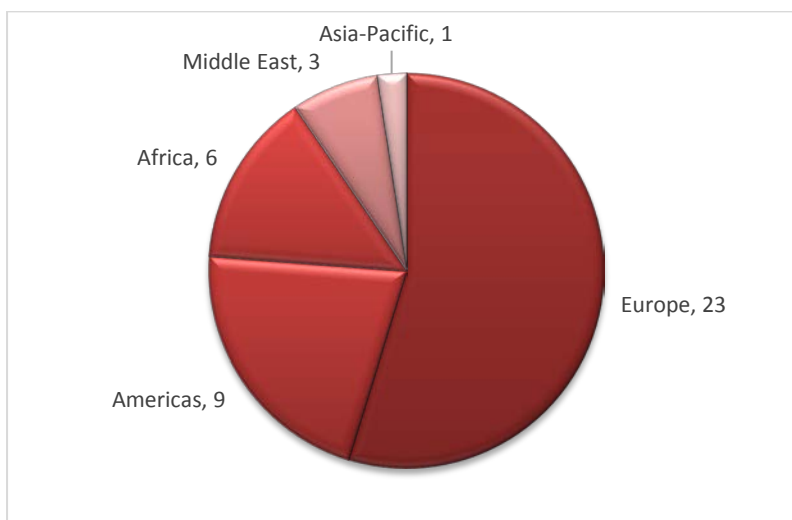
[2] This article is prepared based on the result of the survey conducted by the Central Banking in cooperation with the author and BearingPoint. Detailed results of the survey can be found in the Central Banking Big Data Focus Report 2016.

Staff size

Pie chart – Staff size:
- 2,000-4,999, 2
- 1,000-1,999, 7
- <500, 10
- 500-999, 13

Department

Pie chart – Department:
- Information Technology, 3
- Other, 7
- Statistcs, 32

Central banks have an active interest in big data, particularly improving processing technology, adapting institutional strategies and increasing staff awareness of the area. Just over half of respondents said they had been giving the area active consideration in the past 12 months. This group of 23 central banks was drawn from around the world, though, reflecting the make-up of participants, European central banks figured prominently.

Although big data was not included in its strategic plan for 2016-2021, a respondent from the Caribbean said it would move up on the list of priorities in the future, and, as a result, they have introduced training initiatives.

Has your central bank given any active consideration to big data in the past 12 months?
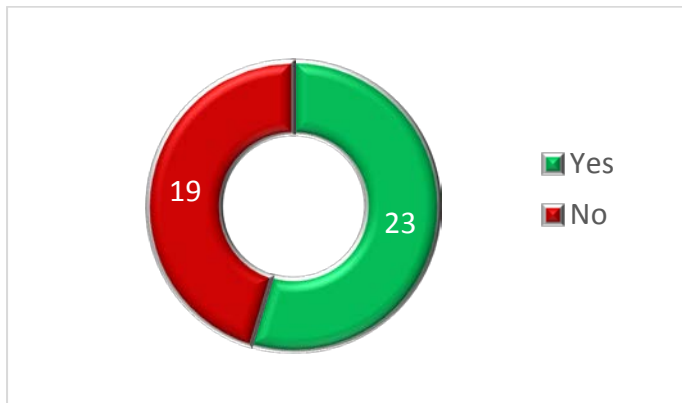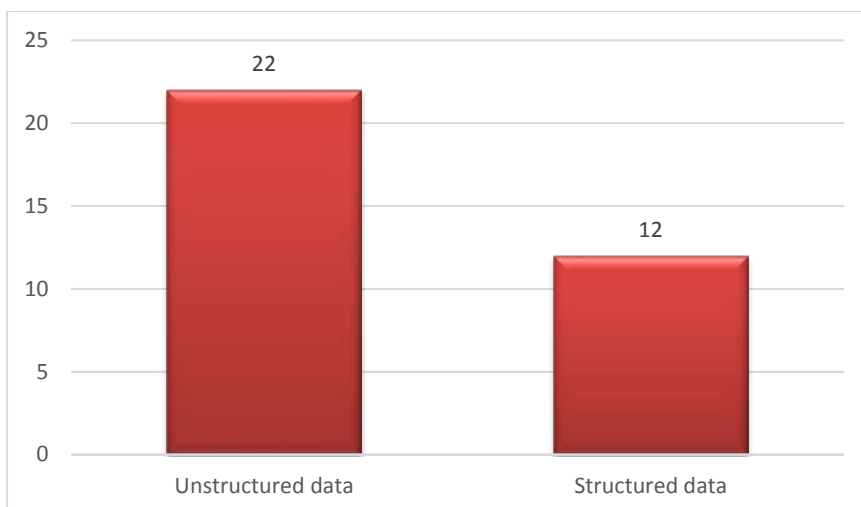

Which of the following types of data would you classify as big data?
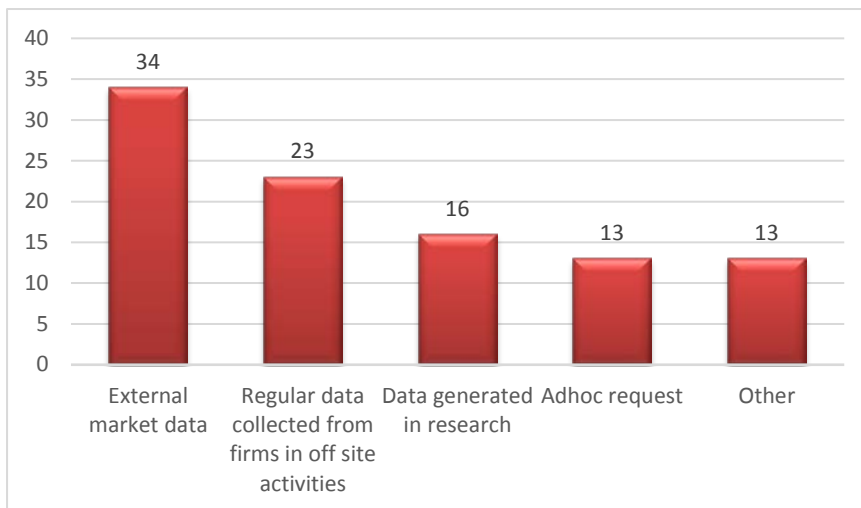
All respondents took part in this question: most gave multiple answers.

Which statement, in your view, represents the most accurate definition of big data?
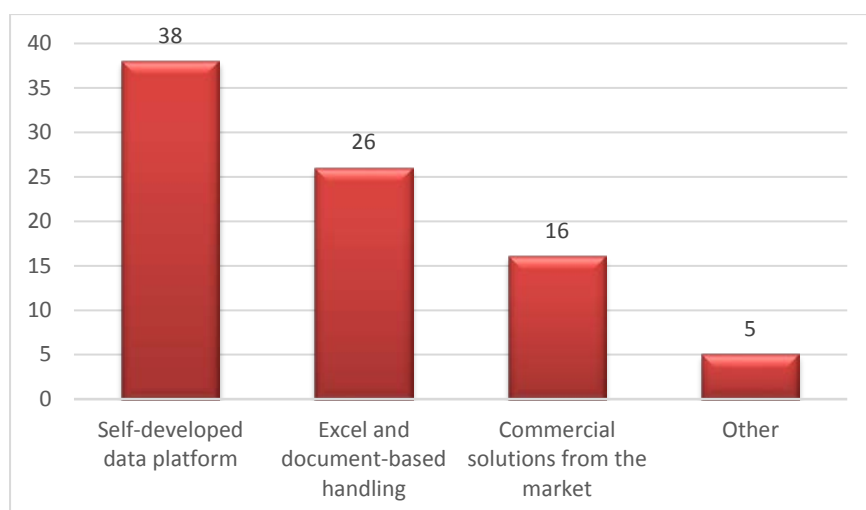
Forty-two central banks answered this question; eight of whom checked both "structured" and "unstructured".

While there is no single agreed definition of big data, industry has gravitated towards viewing it as large volumes of data, both structured and unstructured, which a standard desktop computer cannot handle.[3] Central banks typically see big data as unstructured data that is sourced externally, though this view was not universally held.

Just under two-thirds of respondents believe big data should be defined as unstructured data sets. In a comment, a European central banker said: "'Unstructured' seems to be a more accurate definition of big data because at least it requires unstructured data processing, which is more challenging."

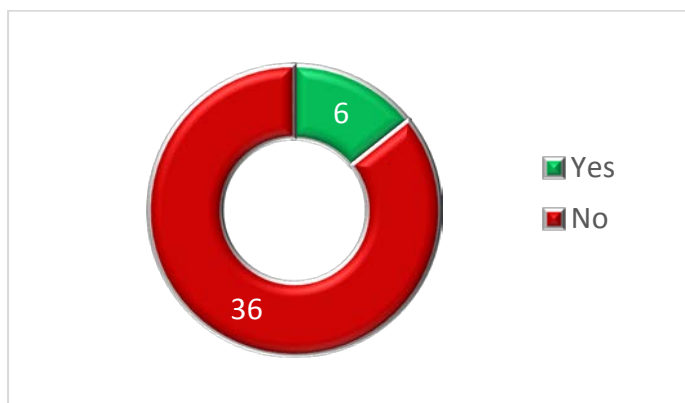**How does your central bank deal with regulatory data collection?**

All respondents gave at least one answer.

As central banks have expanded the depth and breadth of their involvement in financial stability policy-making since the financial crisis, so the importance of collecting accurate, complete and timely data from institutions has increased. Overwhelmingly, central banks develop their own data platforms to handle regulatory data collection. This was the view of 38, or 90%, of the 42 respondents.

Commercial solutions from the market provide a viable alternative for central banks, and this option was chosen by nearly 40% of respondents. This was typically used in conjunction with self-developed platforms, as was the case for three-quarters of this group.

---

[3] For examples of different definitions, please see Bholat, D. (2015) "Big data and central banks" http://www.bankofengland.co.uk/publications/Documents/quarterlybulletin/2015/q108.pdf and Hinge, D. (2015) "The big data revolution and central banks" http://www.centralbanking.com/central-banking-journal/feature/2434258/the-big-data-revolution-and-central-banking.

Has the approach to data collection changed in the past 12 months?

Arrangements for regulatory data collection display a high degree of continuity. Thirty-six respondents said their approach to regulatory data collection was unchanged in the past 12 months. Few respondents volunteered a comment explaining their view, though a handful indicated a transition is underway. A statistician from the Caribbean said their regulatory data collection approach is expected to change over the next six months. Similarly, a respondent from the Americas said: "We are in the process of creating better foundations for policy decisions and have therefore changed strategies for managing data and statistics." One European respondent noted they are reviewing future plans, while another said their central bank is implementing software for both the banking supervision and statistics departments, demonstrating the collaborative nature of big data projects.

Six central banks have changed how they deal with regulatory data collection in the past 12 months. This group was dominated by central banks based in Europe with one adding that it was necessary to change their data collection process "for technological and business reasons."

Which best represents your central bank's view of big data?
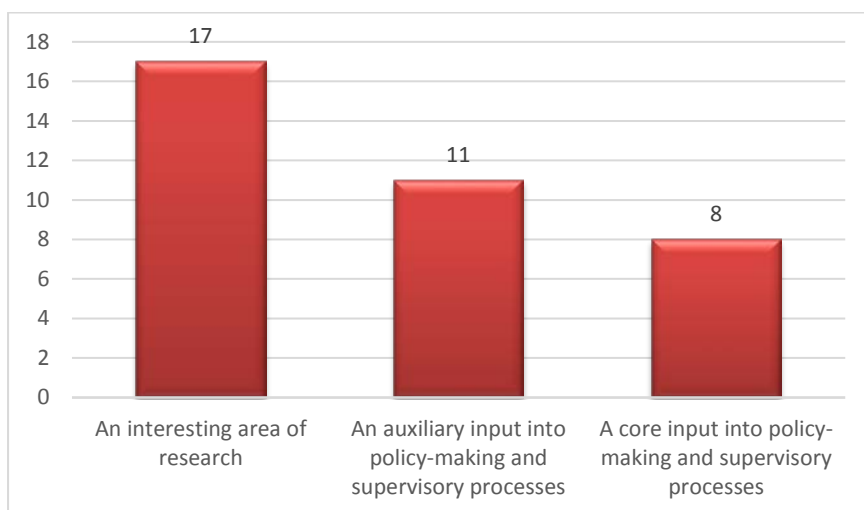
All respondents answered this question: six gave multiple answers.

Big data is predominantly regarded as useful for research in central banks, but significant minorities see immediate involvement in policy-making, or scope for this. Nearly 50% of respondents chose "an interesting area of research" as the best match for how their central bank views big data. A European respondent echoed several others, saying: "the central bank's view of big data is going to change over time. Right now, it is an active area of research". A respondent from an advanced economy implied big data would influence policy: "the optimisation of data is of key importance in driving decision-making."
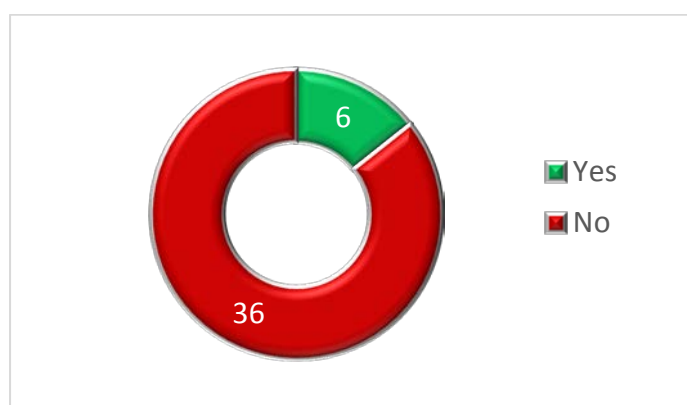
The second most popular choice was 'an auxiliary input'. One European respondent said they could see the potential "in the future" of big data as an auxiliary input in policy and supervision. A central banker from a developing economy said its statistics department manages a handful of micro-databases that are useful for cross-checking data:

The use of microdata also brings flexibility to data management in a way that we can readily adjust and satisfy ad hoc requests, in some cases tailor-made to our customers' needs. Finally, the establishment of protocols with other institutions gives us access to external and complementary information to our own sources, which is one of the primary keys to ensure data quality.

Eight central banks indicated big data was a core input into policy-making and supervisory processes. This included the largest central bank that participated in this survey, as well as two respondents from the Middle East. One commented:

An efficient data management is obviously needed in the course of all missions for which the central bank (including the supervisory authority) is responsible. Big data techniques are useful in that respect.

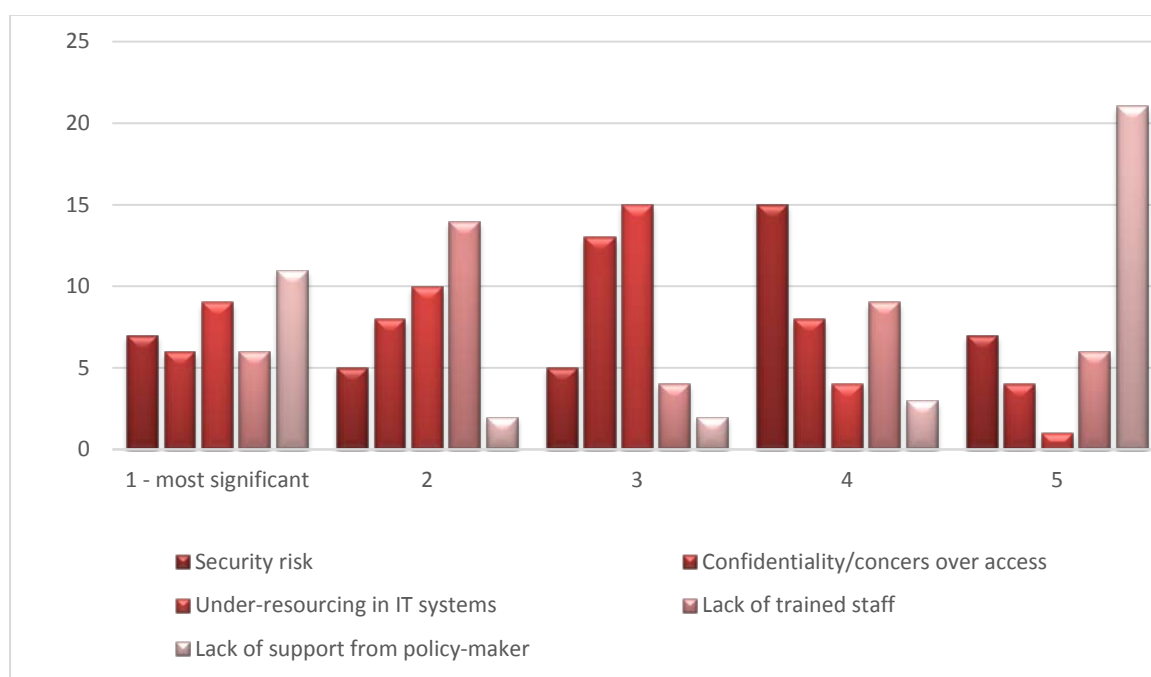**Has your central bank's view changed in the past 12 months?**

Views of big data's role in a central bank are largely unchanged. Over 85% said their view has not changed in the past 12 months. "We expect this to change in the medium term," observed a European central banker, who said they have dedicated a team to this work and therefore their view may change "as a consequence of our research conclusions".

The six central banks that have made a change in the past 12 months highlighted how big data was having a catalytic effect on their institution. A statistician from Europe commented that data management is part of the central bank's new strategic plan. Another European central bank is advancing their work in this area, with dedicated teams within divisions outside of the statistics department:

There has been a change over the past 12-24 months which has seen the creation of dedicated analytics teams within supervisory divisions and a move towards creating a more data-centric organisation, which is reflected in organisational level objectives.

An officer from the Americas looked to the importance of an information strategy for the future development of big data: "To ensure that information gathered is handled in an appropriate manner, a vision for information supply is therefore required, along with an accompanying strategy which guides how data is required and processed."

Which in your view present the greatest roadblocks or challenges to increased use of data sets in your central bank? (Please rank the following 1-5 with 1 being the most significant.)
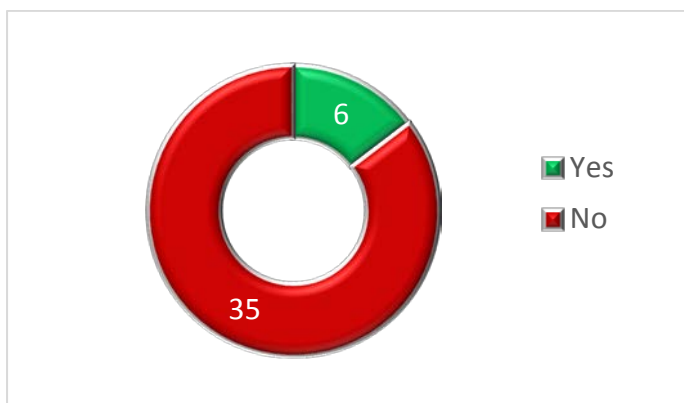

Three respondents did not reply.

Lack of support from policy-makers was most commonly identified as the greatest challenge and, intriguingly, also received the most votes as the least challenging. Twenty-eight per cent of respondents – seven Europeans, and four from the Americas – scored this as the most significant hindrance. "There are no problems with security, but there is not enough support," one European said. Conversely, 21, or 54% of, respondents saw a lack of support from policy-makers as the least significant challenge. This group was largely made up of developing and emerging market economies.

Concerns over skillsets figure prominently in central bankers' thinking. Just over half of respondents placed a lack of trained staff as the first or second most significant challenge. This group contained a significant number of central banks from advanced economies. A respondent from Europe noted: "Specific skills are needed for an efficient management of large data sets, both in IT and statistical departments." One respondent sounded a despondent note: "We expect additional human capital costs to be higher than acceptable compared to the possible benefits from using big data."
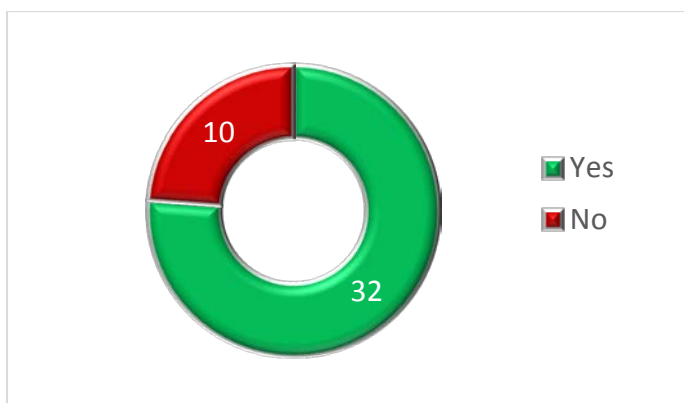
Does your central bank have a single allocated budget for the handling of data (including big data)?

One respondent did not reply.

The vast majority, 85% of respondents, said their central bank did not have a single allocated budget for data. This group was largely made up of central bankers from advanced and developing economies. In comments, respondents typically attributed this to budgeting being divided on departmental or project-specific bases rather than by resource, function or output. In this way, data was often included as part of the technology budget. One central banker commented: "several entities are involved in the handling of data."

Does your central bank have a shared internal platform to enable different areas of the central bank to access data resources?

Over three-quarters of respondents indicated they had a shared internal platform, typically in the form of reporting frameworks and data warehouses. A respondent from a European central bank described the two-pronged approach used in their institution:

We have a shared internal platform for accessing most of our supervision data. For other macroeconomic data we have another platform contributed by the economic department and used by everyone inside the bank.

A respondent from a large central bank explained how they granted access to the data: "The first key element in the data strategy was to allow users with an appropriate business case to view all data relevant to them across the organisation. A technology solution was implemented to allow for this."

If yes, i) can this platform be accessed externally (ie for researchers or dissemination purposes?)
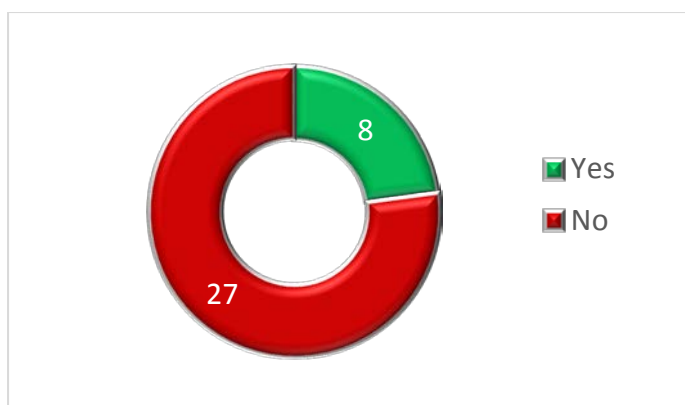
Seven respondents did not reply.

ii) was this shared platform built by the central bank?
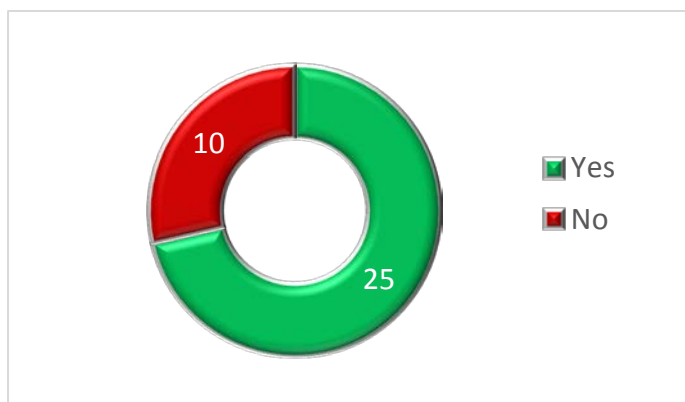
Seven respondents did not reply.

Data-sharing platforms in central banks are typically built internally by the central bank and are not available for external use.

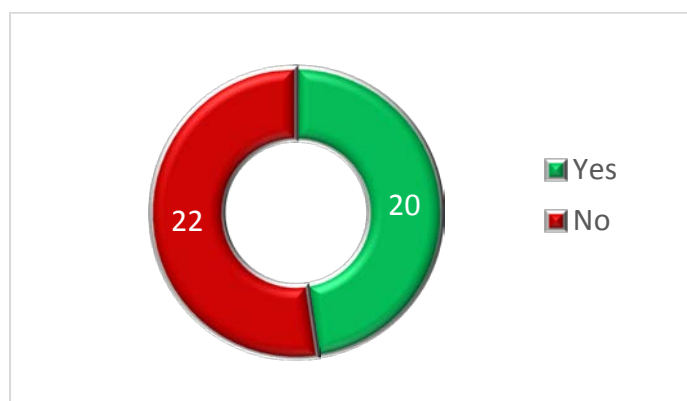The eight central banks that do allow external access to their self-developed platforms consisted of five European central banks, one from the Americas, one from the Middle East and one from Africa. A European central banker said: "the data software is externally accessed by banking supervisors when examining the financial institutions. This platform is restricted for researchers and dissemination purposes." Of the eight central banks that do allow external access, six built their own platform independently.

Several respondents from developed countries, however, said they were establishing platforms that would be available for external use. One central banker said "external access is still under development", while another commented that some parts of the database were being made available to registered users.

In the main, central banks build rather than buy shared platforms: this was the experience of just over 70% of respondents. One European central bank said they used "in-house developments with standard big data infrastructures". Ten central banks indicated they did not build their own platforms. Their reasons ranged from seeking assistance from another central bank in the development to subcontracting the process.

Does your central bank have clear data governance, with defined roles and responsibilities eg, chief data officer, data stewards?
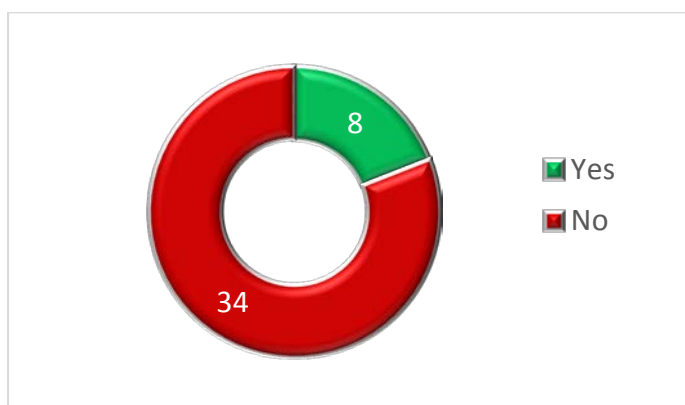

Many central bankers are concerned by data management arrangements, and just over half of respondents said their banks did not have clear data governance. European central banks featured prominently in this group, but there was also a significant number of respondents from the Americas. One said "we have no chief data officer or equivalent".

This is clearly an area of intense activity, however, as central banks are striving to improve data governance frameworks. A central bank from the Caribbean said the framework is "in its infancy" but is "expected to mature" in the coming years. An officer from a central bank in Africa commented a clear data governance structure "is now being put in place". Several central bankers from Europe said it is being discussed, however it had not been formalised, a view typified by this statistician's comment: "Certain relevant people are more aware of their roles and responsibilities as data owners or data stewards, but it has not been fully

formalised and implemented yet." A developed country central banker noted: "We have a CDO and data stewards identified in place, however, more robust governance is being put in place over the coming months."

Twenty respondents were more confident in the arrangements for data governance. One European central banker commented they have "data owners" who can grant others the access to the data. A statistician from the Americas said: "for each information-specialised area, there is a work team." A central bank with over 5,000 staff has committed a whole team to the supervision of big data: "Through the project, we have built a specific organisation in charge of security policy and the high-level supervision. A dedicated committee chaired by the Director of General Statistics including high-level representatives of all business areas is responsible for the high level monitoring of the platform."

Does your central bank have any intra-departmental or divisional bodies, e.g. committees or working groups, dedicated to big data?
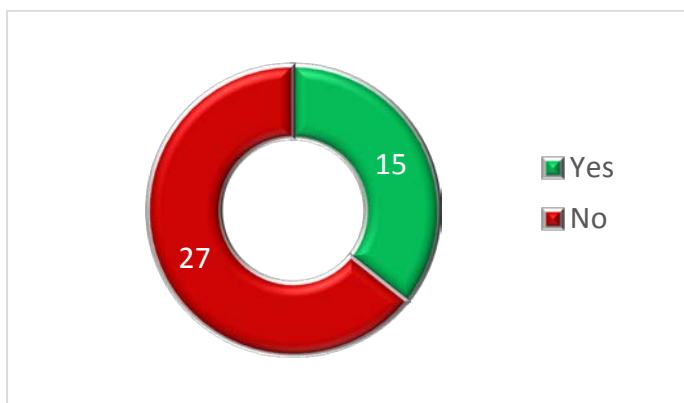

Big data is generally confined to departments or divisions in central banks. Just over 80% of respondents said they do not have any intra-departmental or divisional bodies dedicated to big data. This group featured five African central banks and all nine of the central banks from the Americas, along with two central banks from the Middle East.

Comments made by the 19% of central banks which have intra-departmental or divisional bodies centred around bank-wide data teams and committees dedicated to big data. One central bank commented: "This year we have started an inter-departmental team dedicated to big data issues", while another made reference to big data being discussed within a team of people focused on data in general. One central banker from an advanced economy said big data plays a part throughout the central bank, however it is governed by the statistics department director: "Many projects deal with big data issues: the corresponding steering committees are temporarily chaired by the Director of General Statistics."
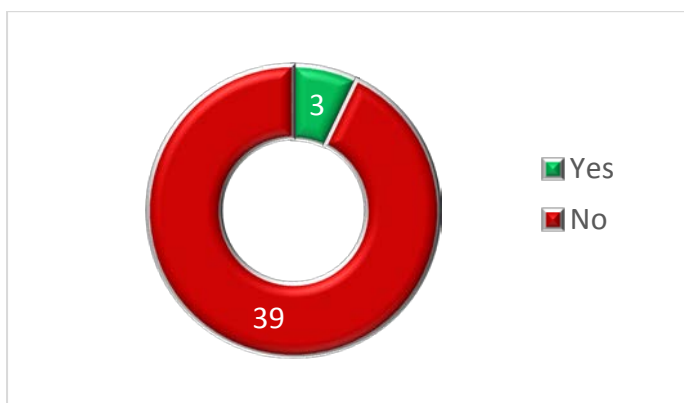
Although 34 central banks indicated they do not have an intra-departmental or divisional bodies at their central banks, three said it is a field they are looking to develop. One such central bank in the Americas commented: "We have some employees working on big data.

Does your central bank use external data providers for big data sets?

Central banks generally source their own big data sets, though a significant minority look elsewhere. Respondents that do not use external data providers were predominantly from emerging-market economies, including 17 central banks from Europe. Those that did use external providers, tended to turn to commercial banks and firms, social media and google blogs, and mobile phone operators. "We employ data from blogs, social media and private websites," said one respondent. A statistician from Europe commented on the techniques used to collect the data from websites: "Currently, we are doing web scraping projects, collecting data from different websites."

Does your central bank outsource any data processing for big data sets?

Overwhelmingly, central banks process their own big data sets and there is no indication of a desire for this to change. More than 90% of respondents answered they do not outsource any data processing. A central bank in the Americas said straightforwardly: "Data processing is run in-house", while one central bank in Europe commented: "The data management and statistical analysis of big data sets is taken on by highly qualified staff in Director General of Statistics with the support of the dedicated team in IT department." Three central banks outsource data processing for big data sets, however they all declined to comment.

Of the 39 central banks who indicated they did not outsource data processing, only three noted this was likely to change in the near future. One central bank in the Americas is

currently hiring in this sector in order to cater for in-house data processing: "The central bank is now recruiting an expert in the field of data architecture in order to develop intra-departmental systems and working routines of data."

In your view which of the following areas stands to benefit most from big data, in practical terms? (Please rank 1-3 with 1 being the most likely to benefit)

Two respondents did not reply.

Respondents saw monetary policy as likely to benefit most from big data, though they said it will have a significant impact on macro-prudential policy as well. As a central banker from an industrial economy explained: "Monetary policy, macro-prudential and micro-prudential policies can benefit from big data. Monetary policy can benefit from better and timelier nowcasts of macroeconomic variables. Macro- and micro-prudential policies might benefit as well."

Interestingly, 90% ranked macro-prudential policy in either first or second place. Those that ranked it in first place typically chose monetary policy as their second choice. Micro-prudential policy was ranked in third place by 58% of respondents but 10 respondents did place it first place. One noted: "Big data applications shine at the most detailed level."

Which area do you consider the priority for investment to increase big data use in your central bank? (Please rank 1-5 with 1 being top priority)
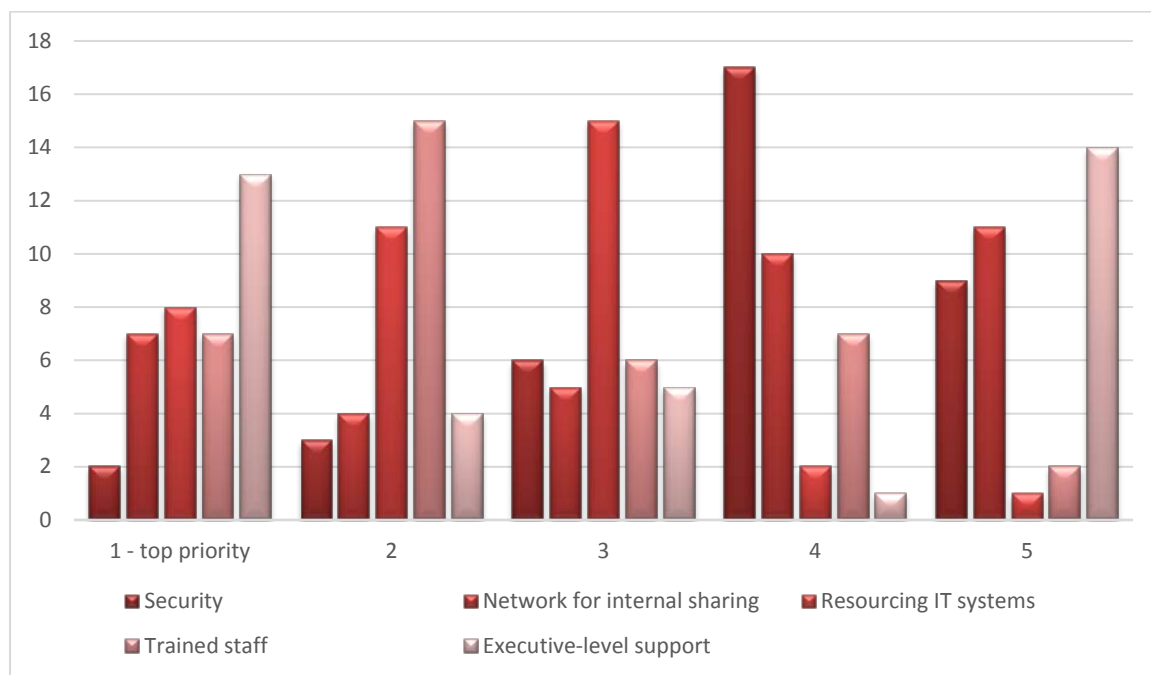
Five respondents did not reply.

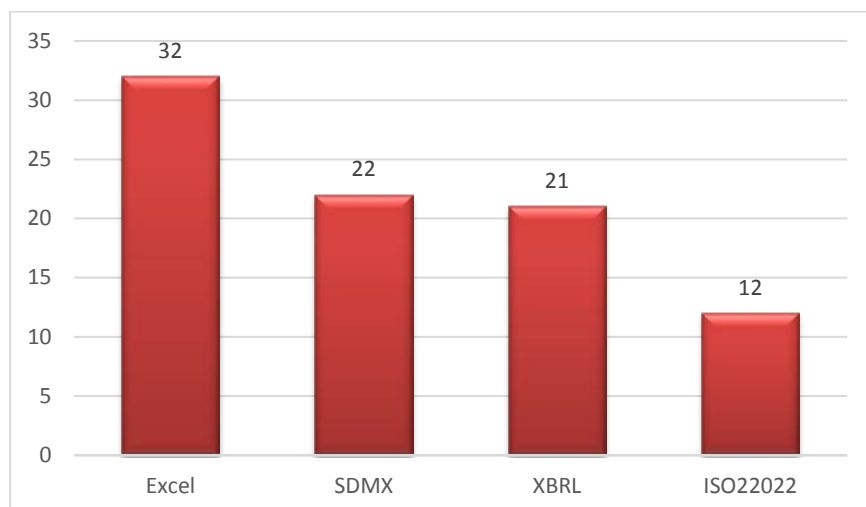Support from the executive-level and policy-makers again divided respondents: 35% saw it as top priority, 38% saw it as the lowest. Those that saw it as a top priority were mainly central banks from Europe and the Americas. A European officer stressed the importance of having a budget: "Obtaining a budget is of fundamental importance in an initiative so therefore is a key priority."

Conversely, many saw executive-level support as a minor challenge. A statistician from a small central bank in the Americas explained their focus:

Big data also poses considerable challenges for the central bank, both technically and methodologically. In addition, new considerations need to be made in terms of strategy, organisation, skills, budgets and risks.

Sixty per cent of respondents ranked trained staff in either first or second place. This group included half a dozen emerging markets. However, these central banks declined to comment. Resourcing IT systems was most commonly ranked in third place. One European central bank commented: "Ensuring that IT infrastructure can handle additional demands is a key component of where this investment will be required." Network for internal sharing was largely ranked in fourth and fifth place by respondents, as was security.

Which of the following standards does your central bank use, or plan to use, for dealing with data exchange and collections?
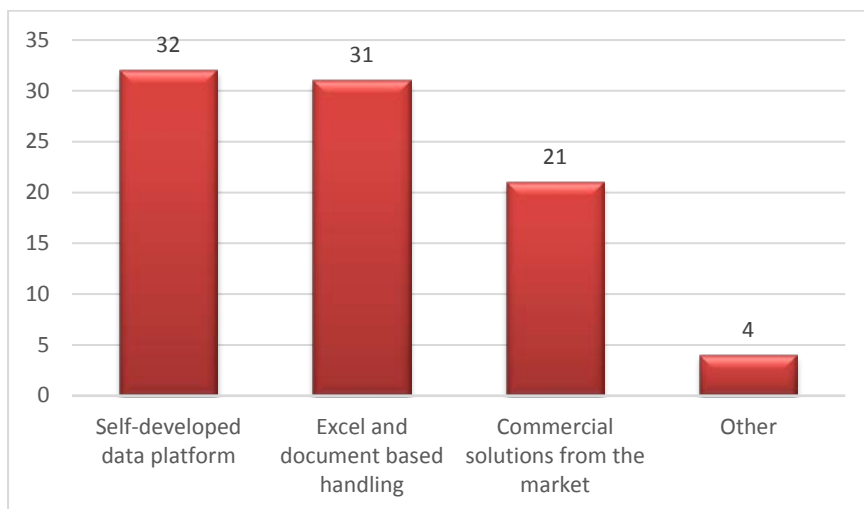

Three respondents did not reply. Respondents checked multiple answers.

Excel is the most popular standard for central banks when dealing with data exchange and collection but it is typically used in conjunction with another solution. Of the 39 respondents, 82% use Excel, and around 40% of those that use Excel use it exclusively. Of the remainder, most used either SDMX or XBRL as well, and three-quarters use both. "The central bank collects data from financial institutions using the XML format combined with Excel format", noted one statistician from a developing economy. An officer from a developed country commented: "XML with structured data prescribed by the central bank is also used for data collection and data exchange".

SDMX was the second most popular standard choice for the majority of developed, European respondents. One central banker commented: "SDMX is used for data exchange of statistical data, primarily in data exchange application with ECB." The statistics department of an African central bank is working towards implementing this standard: "SDMX is an ongoing project."

The standards of XBRL and ISO20022 are used typically for specific functions: supervision and payments, respectively. Twenty-one respondents indicated that they use SDMX at their central bank. A central bank from an advanced economy noted: "XBRL is only used in the transactional system component of RIAD application (Register of Institutions and Affiliates Database)." Although the least popular answer, ISO20022 was likewise described by respondents as useful for specific functions. One European officer noted: "Currently we are using the three standards above-mentioned: SDMX and XBRL are widely used, whereas ISO20022 is only being used in very specific exchanges". An African central bank commented it is under consideration: "ISO20022 is being considered for payments."

**How does your central bank deal with regulatory data analytics?**

One respondent did not reply.

Central banks have developed their own data platforms to deal with regulatory data analytics, but this is often used in conjunction with Excel and document-based handling. This was the answer from over three-quarters of respondents, the majority of which were from developed countries. Only four of the 32 above use purely self-developed regulatory data analytics while the remaining 28 combine it with another standard. Nearly all combined this with Excel and document-based handling and just under half reported a commercial element to their regulatory data analytics workflow.

Commercial solutions from the market proved popular among respondents, with over half choosing this option. Interestingly, more central banks use commercial solutions from the market for regulatory data analytics than they do for regulatory data collection, with only 16 central banks.

As of January 2016 Globally Systemically Important Banks (G-SIBS) have been regulated to meet the BCBS 239 Principles, after carrying out self-assessments, in respect of data collection, data aggregation and dissemination capabilities. Do you think it would be useful for central banks to self-assess using an adapted version of these principles?
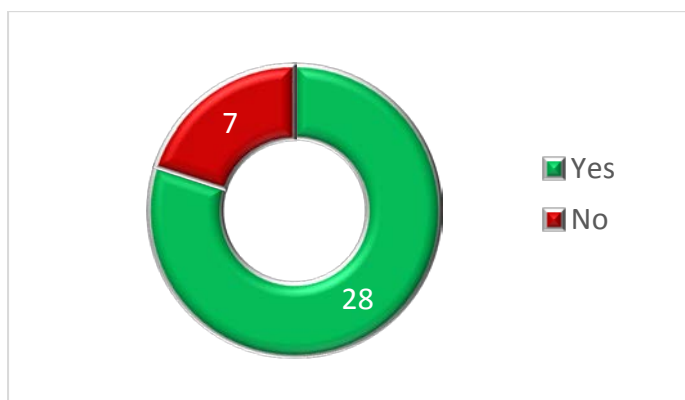
Seven respondents did not reply.

Central bankers largely welcome the idea of self-assessment using an adapted version of the BCBS 239 Principles for Globally Systemically Important Banks (G-SIBs).[4] Of the 35 respondents, 80% said such self-assessment would be useful.

As databases increase in size and complexity, there is naturally a concern that they are policed and governed properly. Furthermore, one central banker noted existing systems were under pressure: "Due to increasingly large and complex data that is now challenging traditional database systems."

Several central banks drew attention to the internal processes already in place to regulate big data processing. One European respondent from an advanced economy noted: "There is a benefit that can be accrued from this. It is important to note however that our internal audit function already undertakes full audits in this area." Another European said BCBS 239 was handled by the supervision department:

However, if the objective of these principles is to strengthen banks' risk data aggregation capabilities and internal risk reporting practices, it might be relevant to make a similar self-assessment in central banks.

Conversely, seven central banks disagreed it would be useful for central banks to self-assess based on an adapted set of principles. All were European except one from the Middle East. The BCBS 239 Principles are not applicable to data collection, a European central banker noted:

BCBS 239 addresses issues pertaining to risk data aggregation and reporting capabilities for banks, so as to achieve full compliance with regulatory expectations. The principles thereof focus on a specific use of data and do not take into account other dimensions of data valuation (e.g. statistical analysis for other needs than those underlying the collection of these data).

Summary

Increasing requirements to central banks' data management due to regulatory trends and technological innovations

Data stay a critical factor for central banks. The financial crisis showed that some of the deepest fissures were caused by gaps in data and exposed the need for high quality, comparable and timely data on the global financial network. Since then, policymakers, supervisory authorities and standard-setters across the globe have been collaborating to greater harmonize and standardize regulatory data in financial services. According to a

---

[4] The BCBS 239 Principles were created by the Basel Committee on Banking Supervision and put into place in January 2016. They consist of 14 principles for supervisors to follow when considering data aggregation in the banks they supervise.

recent BearingPoint Institute paper, urgent debate is still needed on how the world's financial services industry could be better and less onerously supervised via a smarter approach to regulatory reporting and data exchange. [5]

Financial supervision and central banks momentary statistics and financial stability functions are vastly driven by data. In the aftermath of the financial crisis, a "regulatory tsunami" flooded the financial services industry. Especially after the adoption of the Basel III framework, regulatory requirements have significantly increased. New regulations such as AnaCredit, BCBS 239, Solvency 2, Dodd Frank or IFRS 9 have posed new challenges to the banking and insurance sector on global, regional and local levels. Moreover, regulations like the EMIR (European Market Infrastructure Regulation), Money Market Statistical Reporting (MMSR), the Markets in Financial Instruments Regulation (MiFIR) and the Securities Financing Transaction Regulation (SFTR) oblige the major Monetary Financial Institutions (MFIs) to report derivatives or money market data on a daily basis.

"Big data" is a common buzzword in this context. Due to new media and technologies, new data sources appeared like e.g. Internet-based data, data from Social Media, but also from official sources and internal public databases such as banking supervisory data[6]. According to a BearingPoint Institute article on big data, the amount of information available in the world increased by a factor of 100 in last 20 years.

However, in the central banking area, while no single agreed definition exists, big data has already been heralded as offering a wide range of central banking applications: from nowcasting to modelling, to early warning systems and systemic risk indicators. For some it opens a new chapter in policymaking. In a recent study, the Institute of International Finance (IIF) stated that "'RegTech', defined as 'the use of new technologies to solve regulatory and compliance requirements more effectively and efficiently' has enormous potential to … improve the quality and efficiency of supervision, and reduce risk in the system."[7]

According to the 2015 IFC report on "Central banks' use of and interest in 'big data'" central banks have a strong interest in big data, but their actual involvement is still limited.[8]

BearingPoint is noticing two significant key trends worth looking at when discussing (big) data management in central banking in respect to financial services: the replacing of form-based collections with granular, micro-level data[9] and the need to go beyond reporting data validation, i.e. to integrate regulatory Key Performance Indicators (KPIs) into the overall

---

[5] "Reforming Regulatory Reporting. From Templates to Cubes.", Dr. Maciej Piechocki, Tim Dabringhausen
[6] IFC report „Central banks' use of and interest in „big data", October 2015, p. 19
[7] Institute of International Finance, "RegTech in Financial Services: Technology Solutions for Compliance and Reporting.", March 2016, p. 3f
[8] IFC report „Central banks' use of and interest in „big data", October 2015, p. 1
[9] IFC Working Paper No. 14, "Big data: The hunt for timely insights and decision certainty", February 2016, p. 15

operational supervisory process. However a number of further developments at central banks are observable. For example from governance perspectives central banks recently started to appoint "chief data officers" and implement harmonised "data strategies". Numbers of central banks are currently rethinking their data infrastructures which today are rather siloed and demonstrating the legacy of the past decades with no central approach to data handling.

Challenges for central banks

Notwithstanding the huge potential big data provides, decision making is now even harder than before, and business need adequate solutions to analyse this data.[10] A crucial point is how to mine all this information from the different sources exhaustively and at reasonable cost. Despite innovative tools and technologies like blockchain, cloud computing and machine-learning, even today plans often fail because the required processing power outweighs the potential returns or computing time is too long.[11]

The specific challenge for central banks in the sense of an effective 360° risk-based supervision is to rapidly access, effectively manage and timely process and analyse the increasing amounts of supervisory, statistical and markets (big) data. Especially the near or real-time access and efficient processing are regarded as critical factors due to limitations in human and IT resources.[12] According to the IIF report, some regulators still use outdated portal solutions and methods, which are inefficient and increase chances of introducing error.[13] The IIF recommends automated secure data transfer mechanisms based on standards like XBRL (eXtensible Business Reporting Language). But even with use of standard such as XBRL or SMDX (Statistical Data and Metadata eXchange) central banks must abandon "paper-" or "document-oriented" world and think of data in integrated and interlinked manner.

Current systems do not meet today's requirements when regulators have to deal with large amounts of data of various kinds - collected from supervised entities for statistical, prudential or stability purposes, provided by information providers or obtained from internal research and analysis. Such data span from granular micro information on single mortgage loans, securities traded and counterparties affected to macro-economic analysis of countries or regions to form-based collections of financial and risk data or ad-hoc supervisory exercises.

Some of this data will remain only in the perimeter of the central bank some will be remitted to other stakeholders such as the European Supervisory Authorities (ESAs), country

---

[10] BearingPoint Institute Issue 002, "Seeing beyond the big (data) picture, p.3-4
[11] Ibid., p. 6
[12] IFC report „Central banks' use of and interest in „big data", October 2015, p. 11
[13] Institute of International Finance, "RegTech in Financial Services: Technology Solutions for Compliance and Reporting.", March 2016, p. 22-23

governments, the International Monetary Fund (IMF) or the Bank for International Settlements (BIS), some will be disseminated to the wider public or research community.

Therefore, it is mission-critical for regulators to

- effectively handle the large amounts of increasingly granular data from various sources, i.e. rethink existing IT system architectures and landscapes
- gain transparency on the status of the reporting agents in the collection and dissemination process
- consider interlinkages between micro and macro data sets in "going beyond the aggregates" from macro and financial stability perspectives
- get a timely overview of relevant micro and macro developments in the financial markets and
- execute reliable trend analyses on KPIs and Key Risk Indicators (KRIs) based on validated collected data

Essentially, it is a question of scalability in various dimensions across the usual value chain or "lifecycle" of processing supervisory data, as investigated in detail in an article published in Banque de France's Financial Stability Review.[14]

The expanding requirements have proved to be a great challenge and cost driver for IT departments of regulators. IT infrastructure and processes have to be optimised in order to collect, process, analyse and disseminate supervisory and statistical data from different sources and in various formats. But process automation and innovative solutions are required to increase quality and efficiency of supervision, to reduce expenditures, operational burdens and time to market for new supervisory requirements.

---

[14] Mark d. Flood, H.V. Jagadish, Louiqa Raschid: "Big data challenges and opportunities in financial stability monitoring", Banque de France Financial Stability Review, No. 20, April 2016.
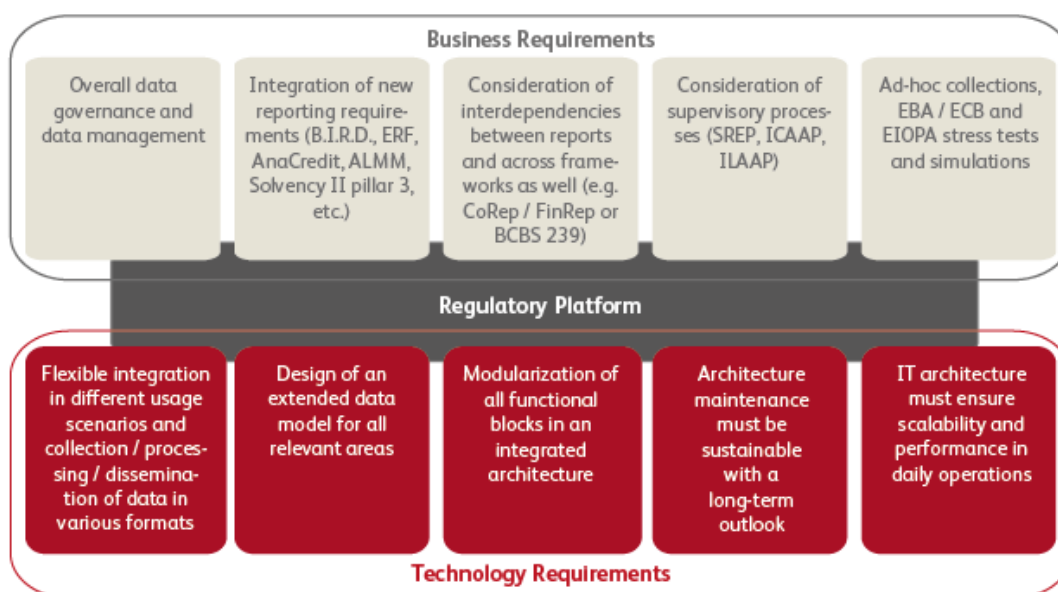
FIGURE 1: REQUIREMENTS FOR FUTURE-ORIENTED REGULATORY PLATFORMS

Innovative approaches – shared utilities, integrated platforms for data management and analytics and Regulatory-as-a-Service (RaaS)

In view of the developments as described before, it is undisputable that it is mission-critical for central banks to reshape their data management and further automate industrialise processes of handling data. Automation helps to minimize risk, reduce errors, and increase transparency and thereby to deliver a better basis for decision-making.

According to a BearingPoint Institute article[15], a new information value chain is needed for reporting which helps to increase efficiency of supervisory processes, minimize risk, allocate resources effectively and improve the basis for decision-making by higher transparency and faster availability of data. We further notice a trend to shared utilities, Regulatory-as-a-Service.

A prominent example is the Austrian solution, where the national central bank, Oesterreichische Nationalbank (OeNB) and the supervised banks joined forces to stepwise replace the template-driven model and use innovative technologies to create a new regulatory value chain. The initiative is based on greater harmonization and integration of data within banks as well as greater integration of the IT systems of the supervisory authority and the supervised entities. The way it works is through a common data model (GMP) developed by central bank in cooperation with Austrian bank and a shared utility, called Austrian Reporting Services GmbH (AuRep), which is co-owned by the largest Austrian

---

[15] BearingPoint Institute, "Reforming regulatory reporting: are we headed toward real-time?", 2015

banking groups. This model allows cost-sharing of compliance as well as standardization of data collection.

AuRep runs on a common platform, which works as the central interface between the banks and the OeNB. Granular bank data sets are captured automatically for supervisors to interrogate in whichever way they want, whilst the banks retain control over their commercially sensitive data, maintaining only the so-called 'passive data interface' on the AuRep platform.

Other regulators are also aware of the limits of the template-based reporting and see the benefits of an input approach with granular datasets. While the Banca d'Italia has been providing such a shared data model named PUMA2 for some decades recently the European Central Bank (ECB) has also launched an initiative to evaluate a European "input approach". The Expert Group on Statistical and Banking Data Dictionary was established to develop a Banks' Integrated Reporting Dictionary (BIRD), which defines a harmonized model for input data as well as rules for the transformation of input data to reporting data. BIRD should be seen as a blueprint for the banks. It forms the conceptual basis of an input approach, i.e., a data model for the organization of the regulatory reporting process within the banks. The approach is similar to the Italian and Austrian model.

Besides harmonized data definitions, new and high-performing supervisory data management platforms are necessary allowing for timely and efficient collection, analysis and sharing of the data.

These platforms could be deployed for instance as a closed solution for the regulator, as an open solution also for firms providing them advanced portal functionality as a service (RaaS or Regulatory-as-a-Service) or even as a shared services platform like in the Austrian case.

With regards to functional scope, new generation platforms should provide functionality for highly automated processing of data and regulatory business intelligence including statistical analysis, monitoring and controlling supervisory Key Risk Indicators (KRIs).

Visions

In the RegTech (Regulatory Technology) developments of recent years, one radical innovation in particular cannot go unnoticed – the Blockchain, and the Distributed Ledger Technology (DLT) that underlies it. Blockchain is rattling conventional finance by transforming business models, connecting new counterparties and eliminating friction. Blockchain offers huge potential cost and time savings in securities and derivatives transactions, especially in clearing, but the fact is that significant work needs to be done before the full benefits of the underlying technology are realised.

For central banks, Blockchain will almost certainly play a role in market infrastructure, although there is not yet a consensus on whether Blockchain is a help or a hindrance to

financial market integration, due to the currently haphazard uptake of numerous different Blockchain technologies. The European Central Bank (ECB), for example, is in the process of assessing the relevance of Blockchain applications for the purposes of improving the region's securities and payments settlement system. However, the ECB has spoken out quite harshly with regard to digital currencies which are underpinned by Blockchain technology, describing them as inherently unstable and presenting a potentially negative impact on the ECB's ability to conduct monetary policy.

At the other end of the spectrum, the Bank of England (a front-runner in this area) has estimated that it could increase GDP by 3% if it introduced a Central Bank Digital Currency (CBDC), as a result of the reduced interest rates, reduced tax rates and lower transaction costs that implementing a digital currency would bring. Such a digital currency is touted as having the potential contribute to the stabilization of the wider economy, as it would give central banks another lever with which to control their currency. This would be particularly effective in times of economic shock, such as Brexit.

In addition to these already significant benefits, a Central Bank Digital Currency could also inject some much needed transparency into the financial system. During the 2008 financial crisis, it became painfully apparent how little visibility was available into the counterparty credit exposure of one major financial institution with regard to others. Eight years down the line, and thousands of pages of regulation later, this requisite level of visibility is still sorely lacking. Blockchain could remediate this opacity, as well as offering protections for privacy.

Following this line of thinking, further visionary uses of Distributed Ledger Technology (DLT) include the delivery of data to regulators, by leveraging so-called Smart Contracts that self-execute in a predefined manner once certain conditions are met. By using Smart Contracts on the Blockchain with specially designed algorithms, stakeholders in the financial system could have the ability to store regulatory information in a secure and immutable form on the shared ledger, and then share it with authorised parties when a specific event occurs. Given the growing multiplicity of data flows, with the ensuing challenges of timeliness, consistency and confidentiality, such an approach could deliver significant value in transaction-based regulatory reporting regimes such as EMIR or MIFID II.

Blockchain and distributed ledger technology (DLT) has a powerful enabling capability that may allow it to become the omnipresent 'soft infrastructure' that strengthens the efficiency, resiliency and accessibility of systems which facilitate global monetary and financial transactions. However, as a technology that is still in its infancy, there are substantial functional, operational, governance and legal aspects which need to be carefully examined before thinking about possible mass adoption. Blockchain and DLT certainly hold great potential, but these fundamental issues must be resolved before the benefits can be fully realized.

IFC-Bank Indonesia Satellite Seminar on *"Big Data"* at the ISI Regional Statistics Conference 2017

Bali, Indonesia, 21 March 2017

Data as a critical factor for central banks[1]

Maciej Piechocki,
BearingPoint / Central Banking

---

[1] This presentation was prepared for the meeting. The views expressed are those of the author and do not necessarily reflect the views of the BIS, the IFC or the central banks and other institutions represented at the meeting.
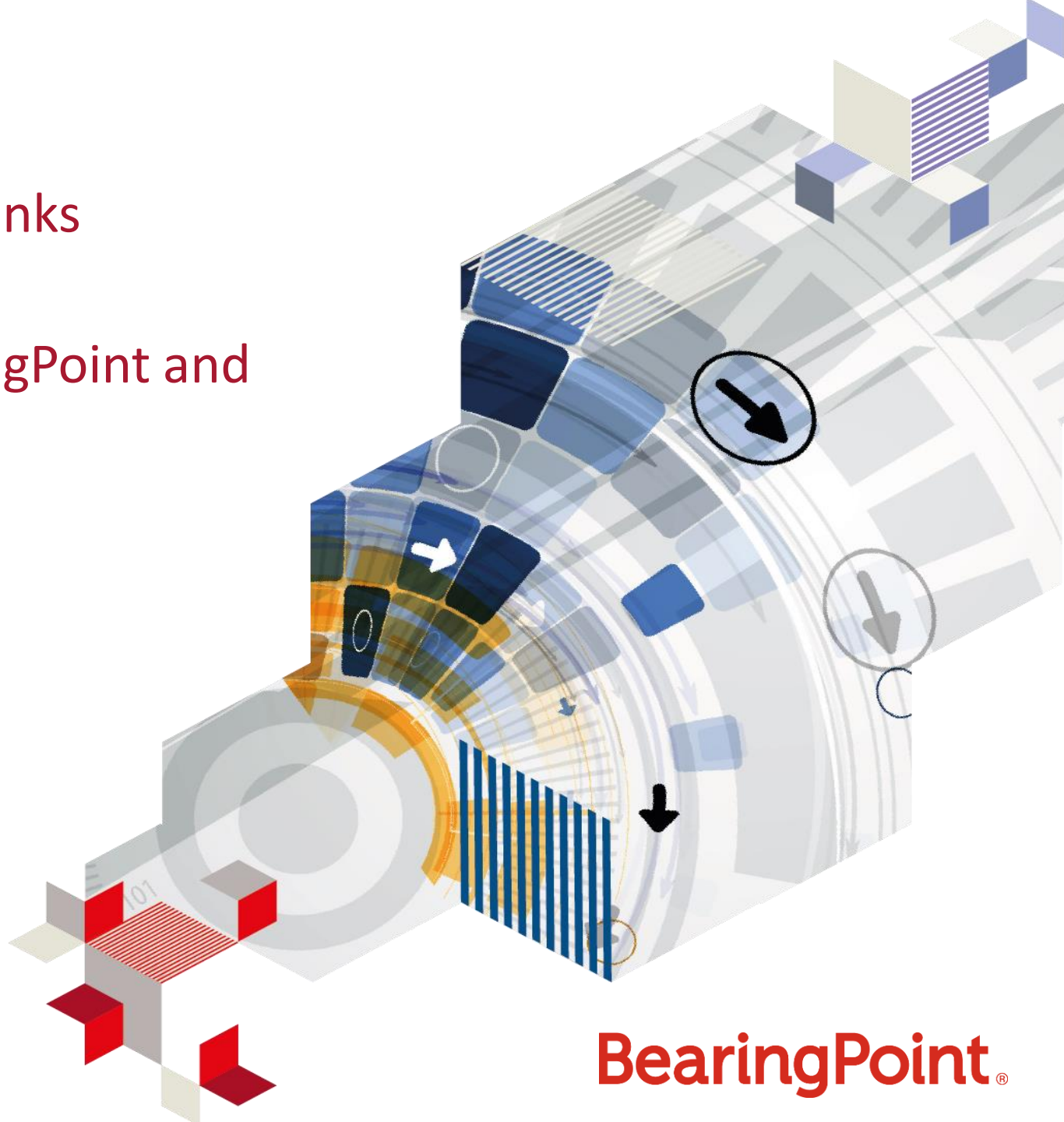
Big Data in Central Banks

Joint survey of BearingPoint and Central Banking

Dr. Maciej Piechocki

Bali, March 21st, 2017

CENTRAL BANKING
AWARDS 2017
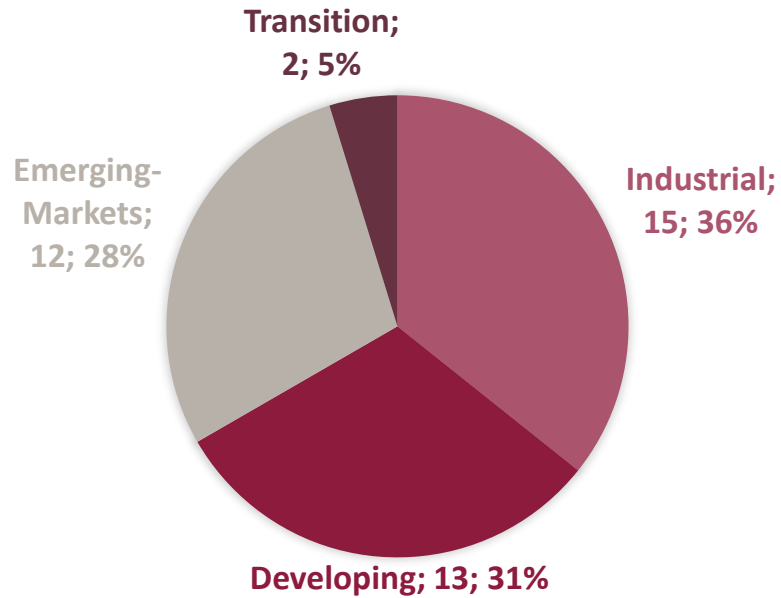CONSULTANCY AND ADVISORY SERVICES
PROVIDER OF THE YEAR (CORE)

CENTRAL BANKING
AWARDS 2016
TECHNOLOGY PROVIDER OF THE YEAR

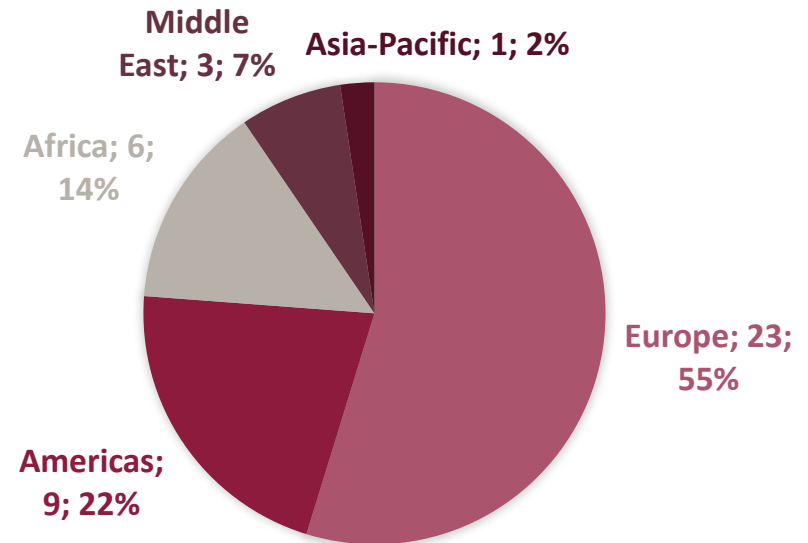BearingPoint®

Economic classification and geographical structure of central banking respondents

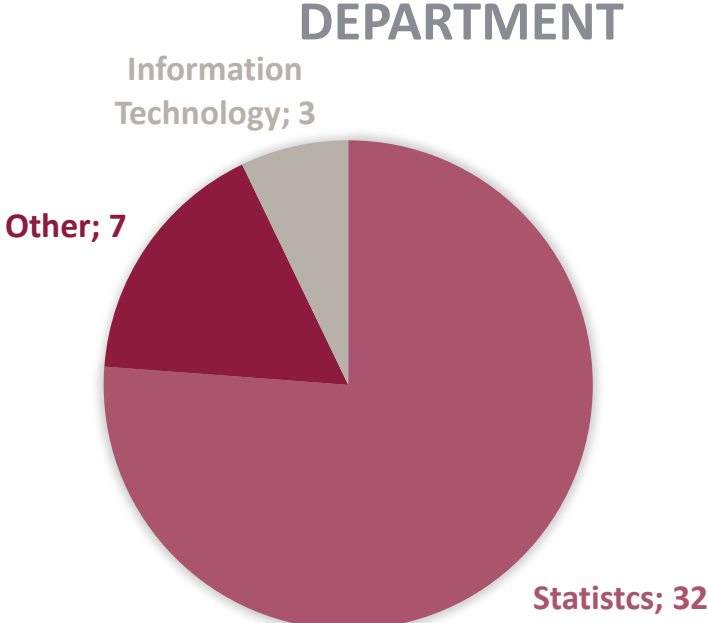**ECONOMIC CLASSIFICATION**

- Transition; 2; 5%
- Emerging-Markets; 12; 28%
- Industrial; 15; 36%
- Developing; 13; 31%

GEOGRAPHY

- Middle East; 3; 7%
- Asia-Pacific; 1; 2%
- Africa; 6; 14%
- Europe; 23; 55%
- Americas; 9; 22%

BearingPoint®

Staff size and departmental structure of central banking respondents

STAFF

- 2,000-4,999; 2
- 1,000-1,999; 7
- <500; 10
- 500-999; 13

DEPARTMENT

- Information Technology; 3
- Other; 7
- Statistcs; 32

BearingPoint.
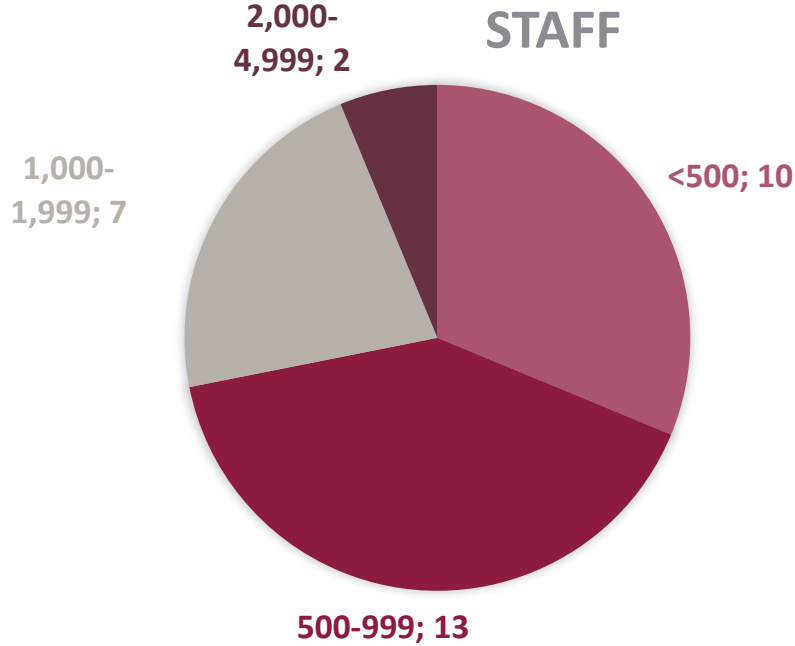
23 out of 42

central banks have an active interest in big data, particularly improving processing technology, adapting institutional strategies and increasing staff awareness

More than

80%
of respondents said an example

of big data was 'external market data'

55%
and over of respondents referred to

'regular data collected from firms in off-site activities'

Big data is predominantly regarded as useful for research in central banks

Monetary policy

likely to benefit most from big data

but also significant impact on macro-prudential policy

BearingPoint.

Over 85% said their view of big data's role in a central bank has not changed in the past 12 months

"We expect this to change in the medium term as a consequence of our research conclusions" *

*a European central banker

Lack of support from policy-makers was most commonly identified as the greatest challenge followed closely by lack of **trained staff**

The vast majority,

85%

of respondents, said their central bank did not have a single allocated budget for data

BearingPoint

Over 3/4

indicated they had a shared internal platform, typically in the form of reporting frameworks and data warehouses

*"We have a shared internal platform for accessing most of our supervision data. For other macroeconomic data we have another platform contributed by the economic department and used by everyone inside the bank." ***

*a European central banker

BearingPoint.

Data-sharing platforms in central banks are typically built **internally** by the central bank and are **not** available for **external use**

Many central bankers are concerned by data management arrangements, and just over half of respondents said their banks have no clear data governance

Just over 80%

of respondents said they do not have any intra-departmental or divisional bodies dedicated to big data

*"This year we have started an inter-departmental team dedicated to big data issues." ***

*a central banker

Top priority for investment to increase big data use in central bank needs support from the

executive-level and policy-makers*

*"Obtaining a budget is of fundamental importance in an initiative so therefore is a key priority." ***

*mainly central banks from Europe and the Americas

**European officer

BearingPoint

38

central banks develop their own data platforms to handle regulatory data collection

For 40% commercial solutions from the market provide a viable alternative

BearingPoint.

Excel is the most popular standard for

82%

of central banks when dealing with data exchange and collection but it is typically used in conjunction with another solution

75% of the central banks have developed their **own data platforms** to deal with regulatory data analytics, but this is often used in conjunction with Excel and document-based handling

BearingPoint.

80% of the central bankers largely welcome the idea of self-assessment using an adapted version of the BCBS 239 Principles for Globally Systemically Important Banks (G-SIBs)

"Due to increasingly large and complex data that is now challenging traditional database systems."*

*a central banker

Business card

BearingPoint ®

Dr. Maciej Piechocki

BearingPoint	T +49 69 13022 6167
Speicherstrasse 1	M +49 152 2286 0072
60327 Frankfurt	
Germany	www.bearingpoint.com

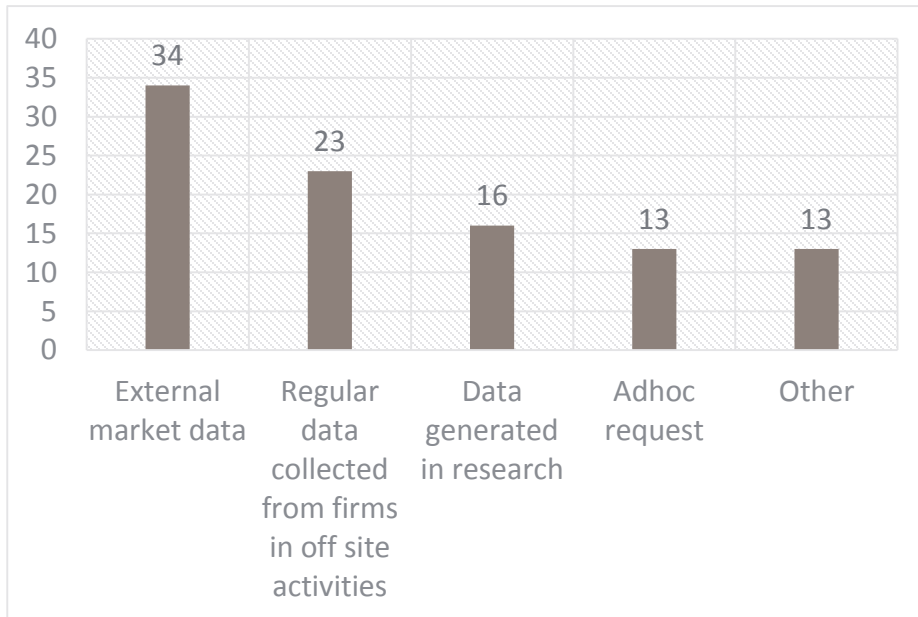maciej.piechocki@bearingpoint.com

BearingPoint ®

BearingPoint ®

Has your central bank given any active consideration to big data in the past 12 months?

- Central banks have an **active interest in big data**, particularly **improving processing technology, adapting institutional strategies and increasing staff awareness** of the area. Just over half of respondents said they had been giving the area active consideration in the past 12 months.

- This group of 23 central banks was drawn from around the world, though, reflecting the make-up of participants, European central banks figured prominently.

- Although big data was not included in its strategic plan for 2016 -2021, a respondent from the Caribbean said it would move up on the list of priorities in the future, and, as a result, they have introduced **training initiatives**.

BearingPoint.

Which statement, in your view, represents the most accurate definition of big data?

Chart data:
- External market data: 34
- Regular data collected from firms in off site activities: 23
- Data generated in research: 16
- Adhoc request: 13
- Other: 13

- More than 80% of respondents said an example of big data was '**external market data**', but only two respondents, both from Europe, checked this option exclusively. Nineteen of the 34 also chose '**regular data collected from firms in off-site activities**'. A statistician from a central bank in Africa included all five classifications in their answer and added "data from ministries, departments and agencies".

- A European central bank ignored the suggested classification and offered the category of "commercial and administrative databases"

All respondents took part in this question:
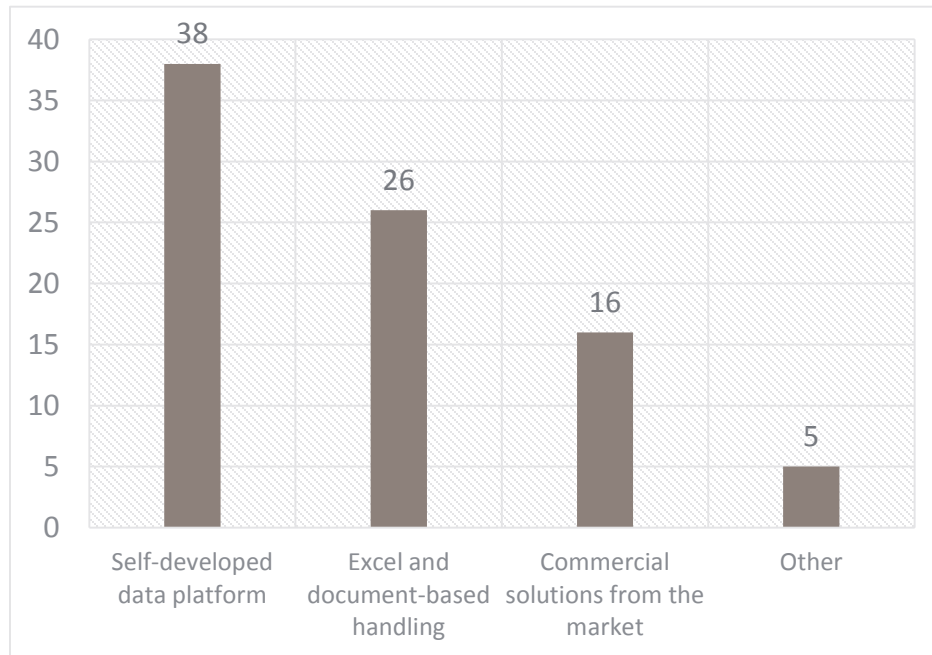most gave multiple answers.

Which of the following types of data would you classify as big data?
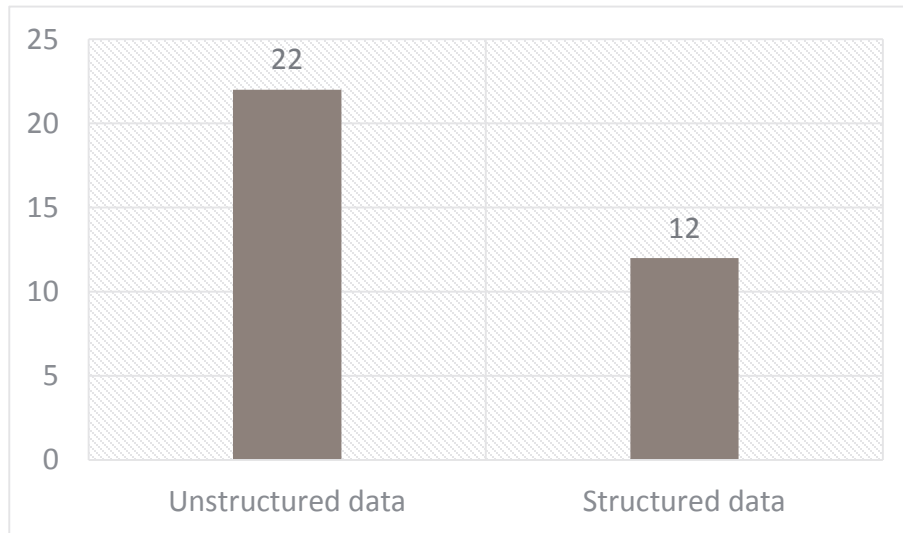
Forty-two central banks answered this question; eight of whom checked both "structured" and "unstructured".

- While there is no single agreed definition of big data, industry has gravitated towards viewing it **as large volumes of data,** both structured and unstructured, which a standard desktop computer cannot handle. Central banks typically see **big data as unstructured data** that is sourced externally, though this view was not universally held.

- Just under two-thirds of respondents believe big data should be defined as unstructured data sets. In a comment, a European central banker said: "'Unstructured' seems to be a more accurate definition of big data because at least it requires unstructured data processing, which is more challenging."
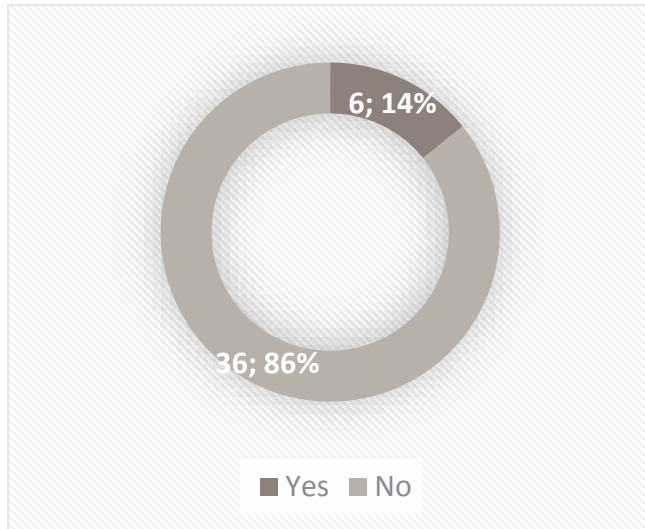
BearingPoint.

How does your central bank deal with regulatory data collection?

Bar chart:
- Self-developed data platform: 38
- Excel and document-based handling: 26
- Commercial solutions from the market: 16
- Other: 5

(Y-axis from 0 to 40)
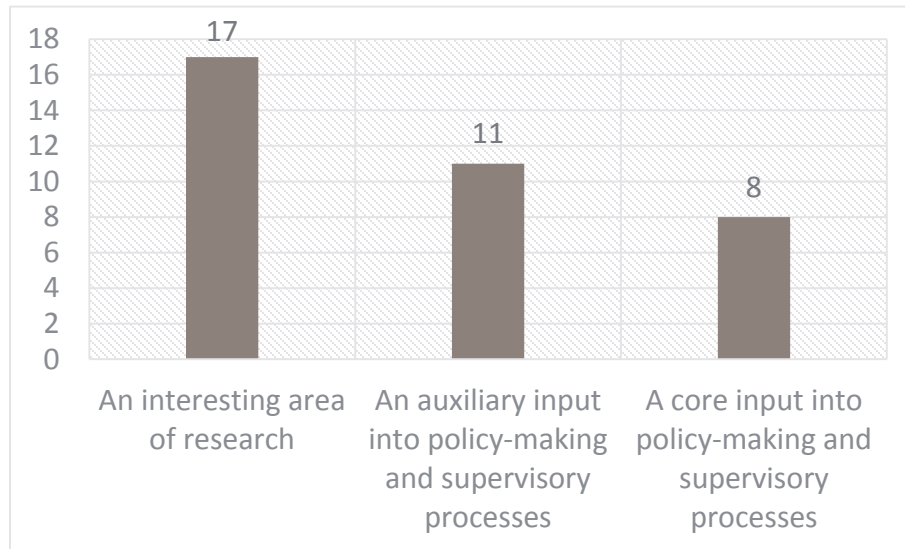
All respondents gave at least one answer.

- As central banks have expanded the depth and breadth of their involvement in financial stability policy-making since the financial crisis, so the importance of collecting accurate, complete and timely data from institutions has increased. Overwhelmingly, central banks **develop their own data platforms to handle regulatory data collection**. This was the view of 38, or 90%, of the 42 respondents.

- **Commercial solutions from the market provide** a viable **alternative** for central banks, and this option was chosen by nearly 40% of respondents. This was typically used in conjunction with self-developed platforms, as was the case for three-quarters of this group.

BearingPoint.

Has the approach to data collection changed in the past 12 months?

6; 14%

36; 86%

■ Yes ■ No

- Arrangements for regulatory data collection display a high degree of continuity. Thirty-six respondents said their approach to regulatory data collection **was unchanged** in the past 12 months. Few respondents volunteered a comment explaining their view, though a handful indicated a transition is underway.

- A statistician from the Caribbean said their regulatory data collection approach is **expected to change over the next six months**. Similarly, a respondent from the Americas said: "We are in the process of creating better foundations for policy decisions and have therefore changed strategies for managing data and statistics." One European respondent noted they are reviewing future plans, while another said their central bank is implementing software for both the banking supervision and statistics departments, demonstrating the collaborative nature of big data projects.

- Six central banks have changed how they deal with regulatory data collection in the past 12 months. This group was dominated by central banks based in Europe with one adding that it was necessary to change their data collection process "for technological and business reasons."
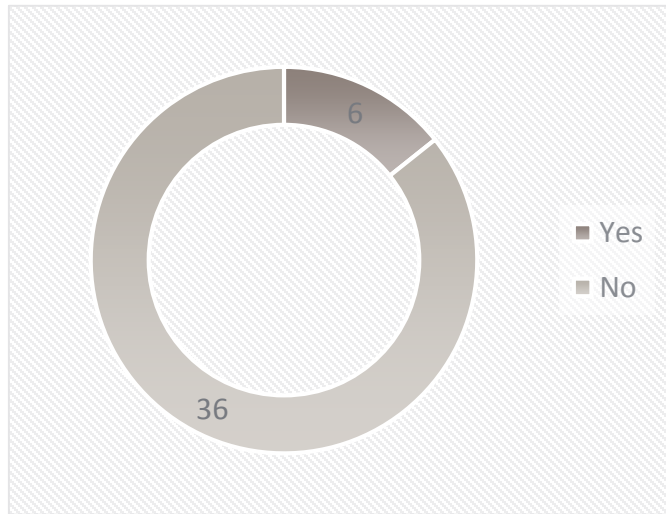
BearingPoint.

Which best represents your central bank's view of big data?

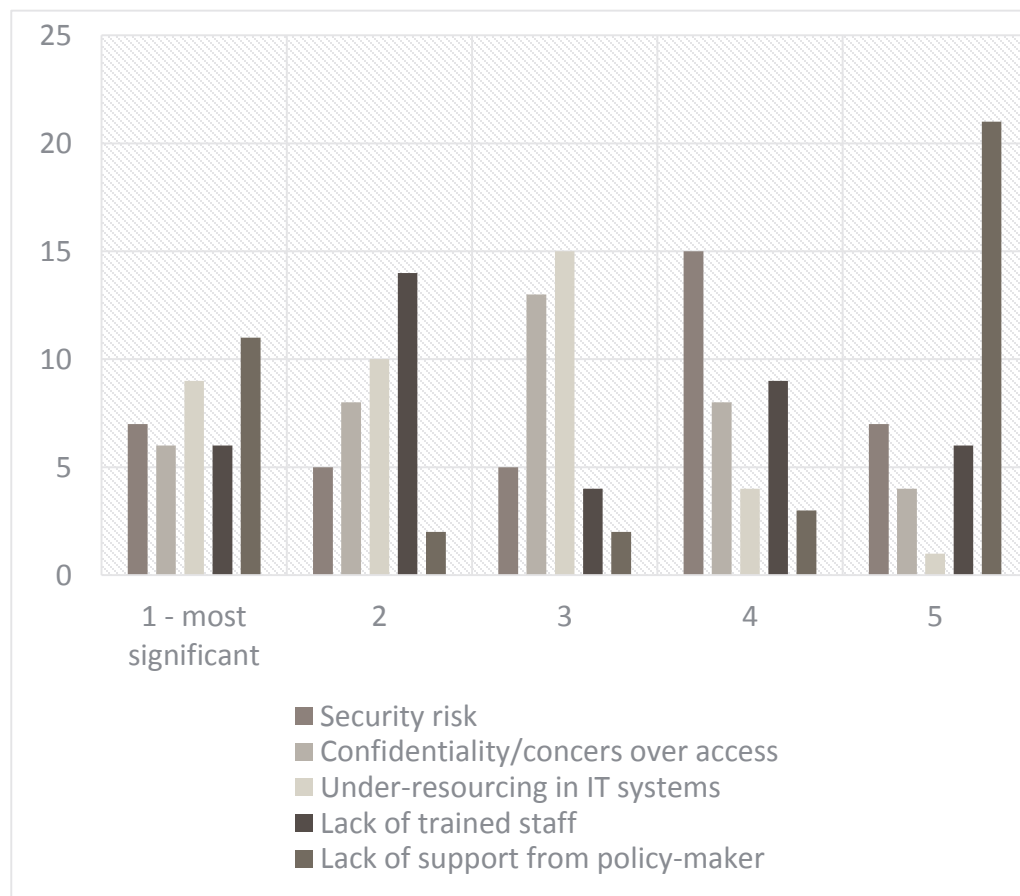All respondents answered this question:
six gave multiple answers.

- Big data is predominantly regarded as **useful for research in central banks**, but significant minorities see immediate involvement in policy-making, or scope for this. Nearly 50% of respondents chose **"an interesting area of research"** as the best match for how their central bank views big data. A European respondent echoed several others, saying: "the central bank's view of big data is going to change over time. Right now, it is an active area of research". A respondent from an advanced economy implied big data would influence policy: "the optimisation of data is of key importance in driving decision-making."

- Eight central banks indicated big data was a core input into policy-making and supervisory processes. This included the largest central bank that participated in this survey, as well as two respondents from the Middle East. One commented:

- An efficient data management is obviously needed in the course of all missions for which the central bank (including the supervisory authority) is responsible. Big data techniques are useful in that respect.

BearingPoint

Has your central bank's view changed in the past 12 months?

6

36

- Yes
- No

- Views of big data's role in a central bank are largely unchanged. Over 85% said their **view has not changed** in the past 12 months.

- "We expect this to change in the medium term," observed a European central banker, who said they have dedicated a team to this work and therefore their view may change "as a consequence of our research conclusions".

- The six central banks that have made a change in the past 12 months **highlighted how big data was having a catalytic effect on their institution**. A statistician from Europe commented that data management is part of the central bank's new strategic plan. Another European central bank is advancing their work in this area, with dedicated teams within divisions outside of the statistics department: "There has been a change over the past 12-24 months which has seen the creation of dedicated analytics teams within supervisory divisions and a move towards creating a more data-centric organisation, which is reflected in organisational level objectives."

- An officer from the Americas looked to the importance of an information strategy for the future development of big data: "To ensure that information gathered is handled in an appropriate manner, a **vision for information supply is therefore required, along with an accompanying strategy which guides how data is required and processed**."

BearingPoint.

Which in your view present the greatest roadblocks or challenges to increased use of data sets in your central bank?

Bar chart legend:
- Security risk
- Confidentiality/concers over access
- Under-resourcing in IT systems
- Lack of trained staff
- Lack of support from policy-maker
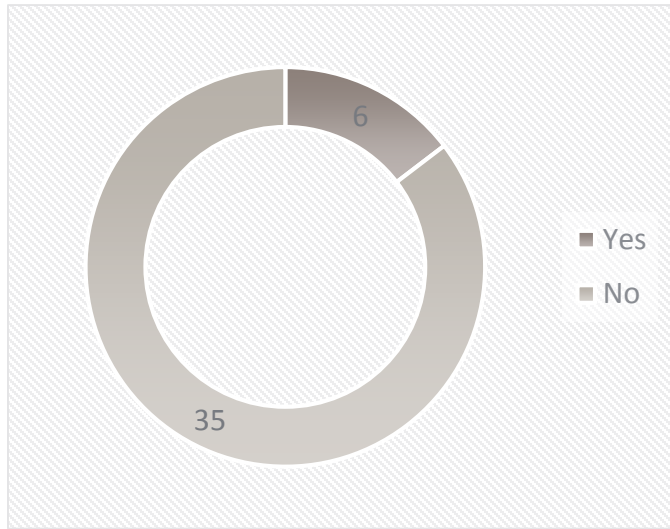
X-axis: 1 - most significant, 2, 3, 4, 5

Votes were cast using a scale of 1-5, where 1 denotes the most significant roadblock, and 5 the least significant.
Three respondents did not reply.

- **Lack of support from policy-makers** was most commonly identified as the greatest challenge and, intriguingly, also received the most votes as the least challenging. Twenty-eight per cent of respondents – seven Europeans, and four from the Americas – scored this as the most significant hindrance. "There are no problems with security, but there is not enough support," one European said. Conversely, 21, or 54% of, respondents saw a lack of support from policy-makers as the least significant challenge. This group was largely made up of developing and emerging market economies.

- Concerns over skillsets figure prominently in central bankers' thinking. Just over half of respondents placed a lack of trained staff as the first or second most significant challenge.
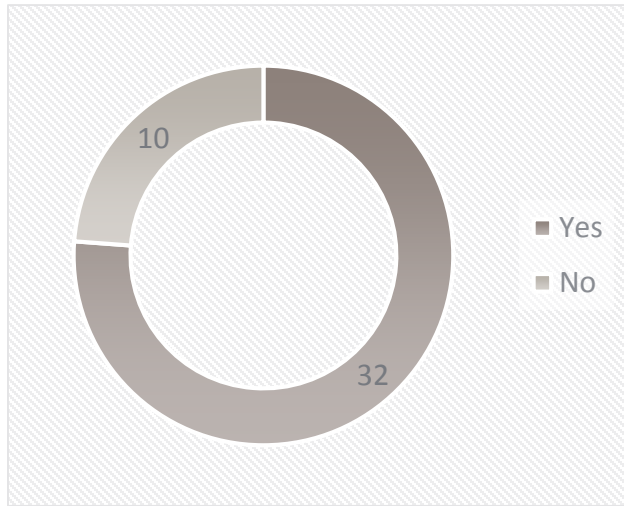
BearingPoint

Does your central bank have a single allocated budget for the handling of data (including big data)?

6

35

■ Yes
■ No

- The vast majority, 85% of respondents, said their central bank **did not have a single allocated budget** for data. This group was largely made up of central bankers from advanced and developing economies.

- In comments, respondents typically attributed this to budgeting being **divided on departmental or project-specific bases r**ather than by resource, function or output. In this way, data was often included as part of the technology budget. One central banker commented: "several entities are involved in the handling of data."

One respondent did not reply.

Does your central bank have a shared internal platform to enable different areas of the central bank to access data resources?
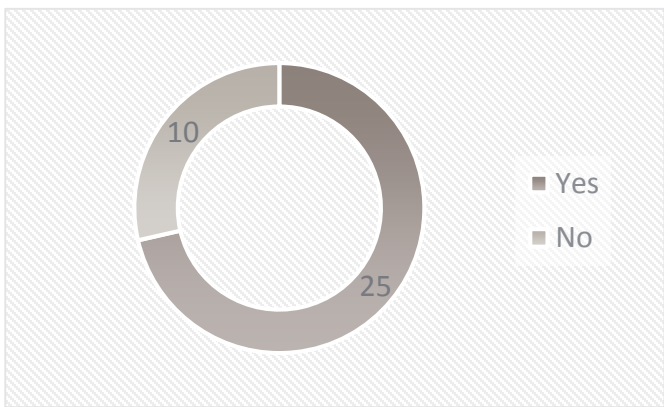
10

Yes

No

32

- Over three-quarters of respondents indicated they **had a shared internal platform**, typically in the form of reporting frameworks and data warehouses. A respondent from a European central bank described the two-pronged approach used in their institution: "We have a shared internal platform for accessing most of our supervision data. For other macroeconomic data we have another platform contributed by the economic department and used by everyone inside the bank."

- A respondent from a large central bank explained how they granted access to the data: "The first key element in the data strategy was to allow users with an appropriate business case to view all data relevant to them across the organisation. A technology solution was implemented to allow for this."

BearingPoint.

Does your central bank have a shared internal platform to enable different areas of the central bank to access data resources?

If yes,
i) can this platform be accessed externally (ie for researchers or dissemination purposes?)
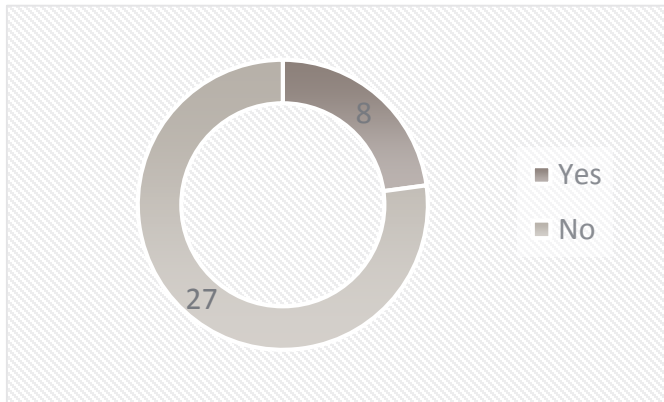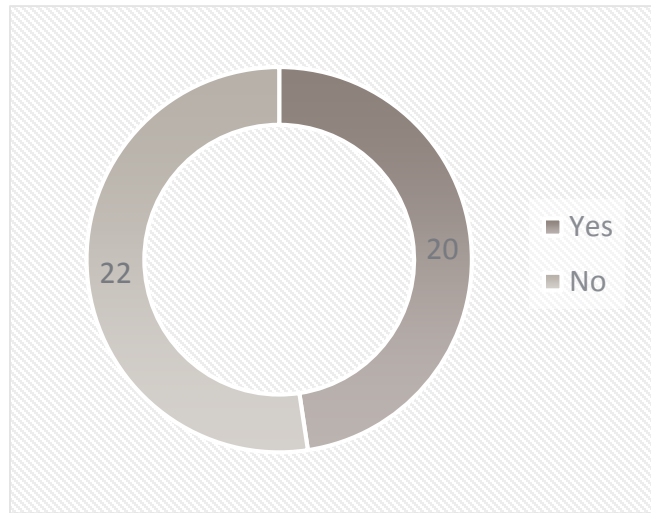

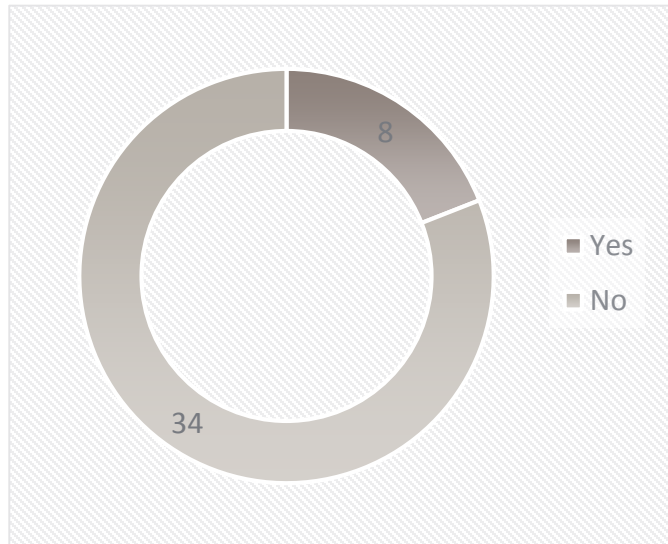**ii) was this shared platform built by the central bank?**

- Data-sharing platforms in central banks are typically built **internally** by the central bank and are not available for external use.

- The eight central banks that do **allow external access** to their self-developed platforms consisted of five European central banks, one from the Americas, one from the Middle East and one from Africa. A European central banker said: "the data software is externally accessed by banking supervisors when examining the financial institutions. This platform is restricted for researchers and dissemination purposes." Of the eight central banks that do allow external access, six built their own platform independently.

- Several respondents from developed countries, however, said they **were establishing platforms that would be available for external use**. One central banker said "external access is still under development", while another commented that some parts of the database were being made available to registered users.

- In the main, central banks **build rather than buy shared platforms**: this was the experience of just over 70% of respondents. One European central bank said they used "in-house developments with standard big data infrastructures". Ten central banks indicated they did not build their own platforms. Their reasons ranged from seeking assistance from another central bank in the development to subcontracting the process.

BearingPoint.

Does your central bank have clear data governance, with defined roles and responsibilities eg, chief data officer, data stewards?
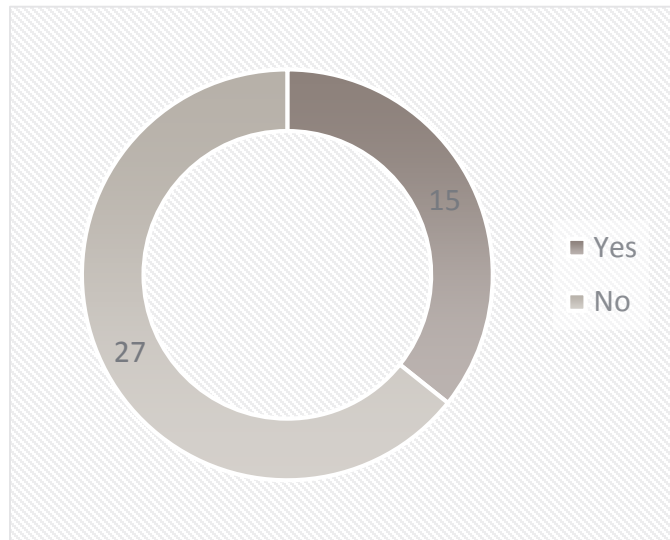

Donut chart:
- Yes: 20
- No: 22

- Many central bankers are **concerned by data management arrangements**, and just over half of respondents said their banks **did not have clear data governance**. European central banks featured prominently in this group, but there was also a significant number of respondents from the Americas. One said "we have no chief data officer or equivalent".

- This is clearly an area of intense activity, however, as central banks are striving to improve data governance frameworks.

- A central bank from the Caribbean said the framework is "in its infancy" but is "expected to mature" in the coming years. An officer from a central bank in Africa commented a clear data governance structure "is now being put in place". Several central bankers from Europe said it is being discussed, however it had not been formalised, a view typified by this statistician's comment: "Certain relevant people are more aware of their roles and responsibilities as data owners or data stewards, but it has not been fully formalised and implemented yet." A developed country central banker noted: "We have a CDO and data stewards identified in place, however, more robust governance is being put in place over the coming months."

BearingPoint®

Does your central bank have any intra-departmental or divisional bodies, e.g. committees or working groups, dedicated to big data?
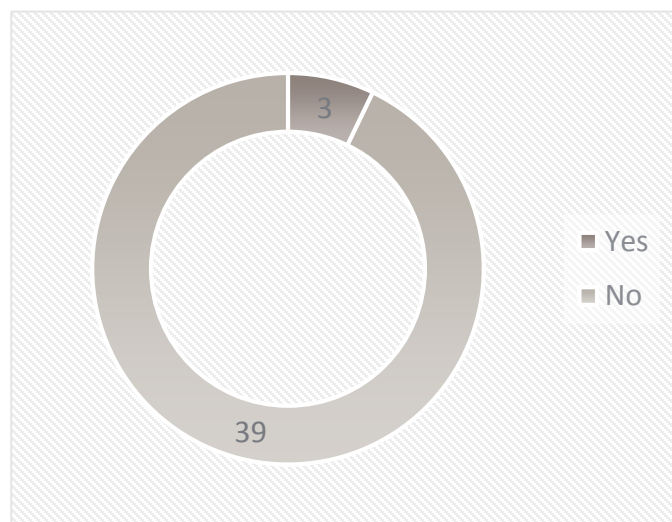

- Big data is **generally confined to departments or divisions** in central banks. Just over 80% of respondents said they **do not have** any intra-departmental or divisional bodies dedicated to big data. This group featured five African central banks and all nine of the central banks from the Americas, along with two central banks from the Middle East.

- Comments made by the 19% of central banks which have intra-departmental or divisional bodies centred around bank-wide data teams and committees dedicated to big data. One central bank commented: "This year we have started an inter-departmental team dedicated to big data issues", while another made reference to big data being discussed within a team of people focused on data in general. One central banker from an advanced economy said big data plays a part throughout the central bank, however it is governed by the statistics department director: "Many projects deal with big data issues: the corresponding steering committees are temporarily chaired by the Director of General Statistics."

- Although 34 central banks indicated they do not have an intra-departmental or divisional bodies at their central banks, three said it is a field they are looking to develop. One such central bank in the Americas commented: "We have some employees working on big data.

Does your central bank use external data providers for big data sets?
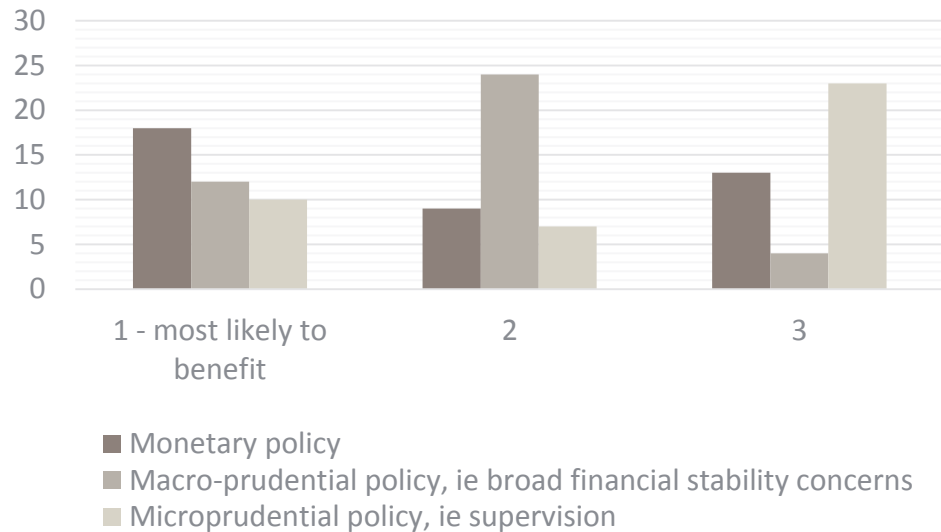

Legend:
- ■ Yes (15)
- ■ No (27)

- Central banks **generally source their own big data sets**, though a significant minority look elsewhere. Respondents that do not use external data providers were predominantly from emerging-market economies, including 17 central banks from Europe.

- Those that did use external providers, tended to turn to commercial banks and firms, social media and google blogs, and mobile phone operators. "We employ data from blogs, social media and private websites," said one respondent. A statistician from Europe commented on the techniques used to collect the data from websites: "Currently, we are doing web scraping projects, collecting data from different websites."

BearingPoint.

Does your central bank outsource any data processing for big data sets?

- Overwhelmingly, **central banks process their own big data sets** and there is **no indication of a desire for this to change**. More than 90% of respondents answered they do not outsource any data processing.

- A central bank in the Americas said straightforwardly: "Data processing is run in-house", while one central bank in Europe commented: "The data management and statistical analysis of big data sets is taken on by highly qualified staff in Director General of Statistics with the support of the dedicated team in IT department." Three central banks outsource data processing for big data sets, however they all declined to comment.

- Of the 39 central banks who indicated they did not outsource data processing, only three noted this was likely to change in the near future. One central bank in the Americas is currently hiring in this sector in order to cater for in-house data processing: "The central bank is now recruiting an expert in the field of data architecture in order to develop intra-departmental systems and working routines of data."
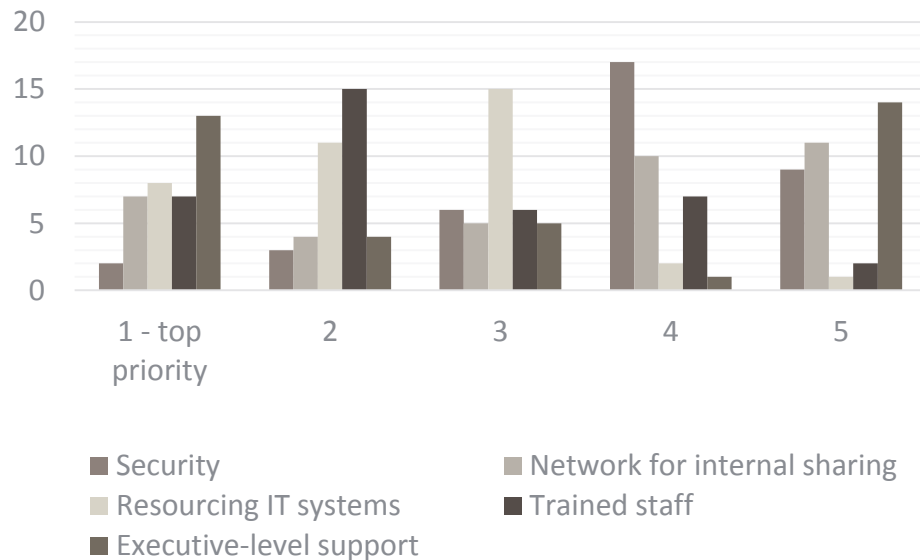
BearingPoint.

In your view which of the following areas stands to benefit most from big data, in practical terms?

Chart legend:
- Monetary policy
- Macro-prudential policy, ie broad financial stability concerns
- Microprudential policy, ie supervision

X-axis categories: 1 - most likely to benefit, 2, 3
Y-axis: 0 to 30

Two respondents did not reply.

- Respondents saw **monetary policy as likely to benefit most from big data,** though they said it will have a significant impact on macro-prudential policy as well. As a central banker from an industrial economy explained: "Monetary policy, macro-prudential and micro-prudential policies can benefit from big data. Monetary policy can benefit from better and timelier nowcasts of macroeconomic variables. Macro- and micro-prudential policies might benefit as well."

- Interestingly, **90% ranked macro-prudential policy in either first or second place**. Those that ranked it in first place typically chose monetary policy as their second choice. Micro-prudential policy was ranked in third place by 58% of respondents but 10 respondents did place it first place.
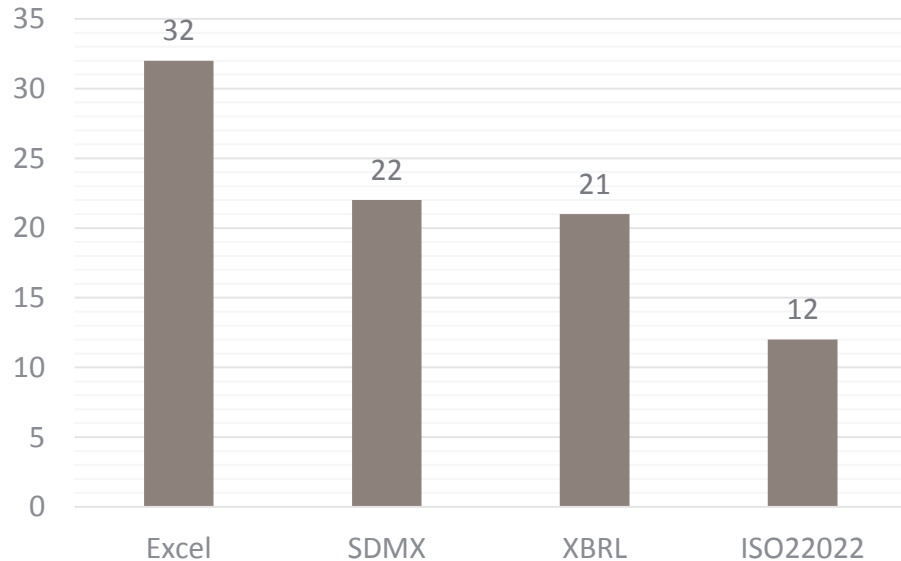
BearingPoint.

Which area do you consider the priority for investment to increase big data use in your central bank?

Five respondents did not reply.

- **Support from the executive-level and** policy-makers again divided respondents: 35% saw it as top priority, 38% saw it as the lowest. Those that saw it as a top priority were mainly central banks from Europe and the Americas. A European officer stressed the importance of having a budget: "Obtaining a budget is of fundamental importance in an initiative so therefore is a key priority."

- Sixty per cent of respondents ranked trained staff in either first or second place. This group included half a dozen emerging markets. However, these central banks declined to comment. Resourcing IT systems was most commonly ranked in third place. One European central bank commented: "Ensuring that IT infrastructure can handle additional demands is a key component of where this investment will be required." Network for internal sharing was largely ranked in fourth and fifth place by respondents, as was security.
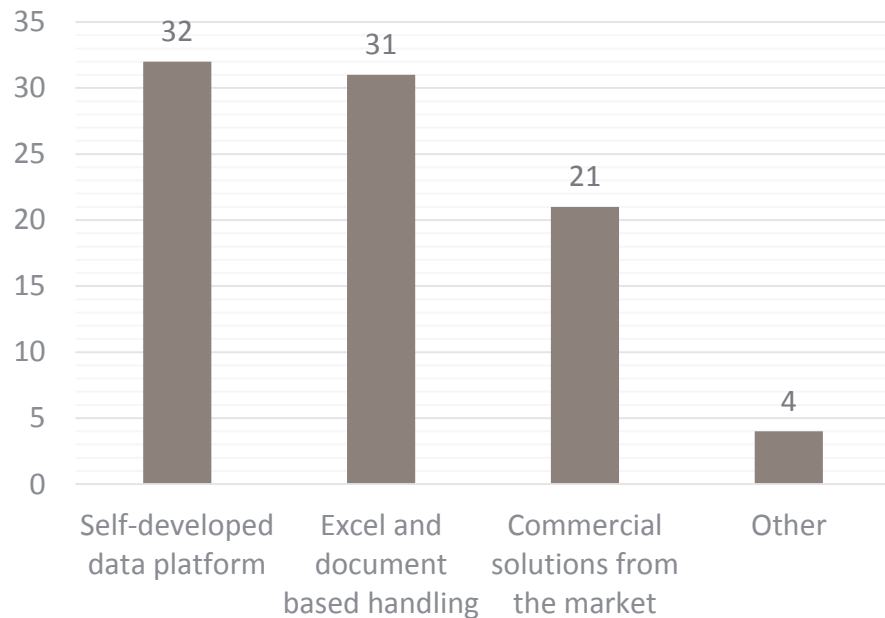
BearingPoint.

Which of the following standards does your central bank use, or plan to use, for dealing with data exchange and collections?

Three respondents did not reply.
Respondents checked multiple answers.

- **Excel is the most popular standard** for central banks when dealing with data exchange and collection but it is typically used in conjunction with another solution. Of the 39 respondents, 82% use Excel, and around 40% of those that use Excel use it exclusively. Of the remainder, most used either SDMX or XBRL as well, and three-quarters use both.

- The standards of XBRL and ISO20022 are used typically for specific functions: supervision and payments, respectively. Twenty-one respondents indicated that they use SDMX at their central bank. A central bank from an advanced economy noted: "XBRL is only used in the transactional system component of RIAD application (Register of Institutions and Affiliates Database)." Although the least popular answer, ISO20022 was likewise described by respondents as useful for specific functions. One European officer noted: "Currently we are using the three standards above-mentioned: SDMX and XBRL are widely used, whereas ISO20022 is only being used in very specific exchanges".
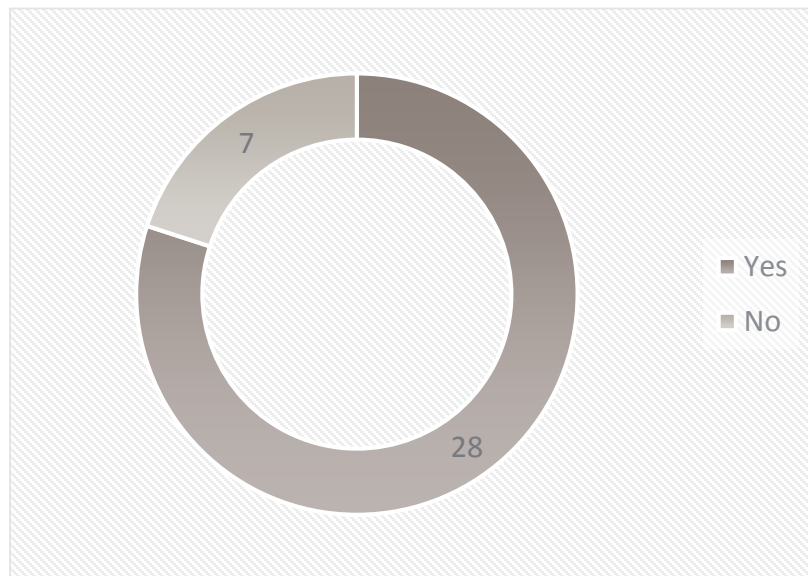
BearingPoint.

How does your central bank deal with regulatory data analytics?

32 — Self-developed data platform
31 — Excel and document based handling
21 — Commercial solutions from the market
4 — Other

One respondents did not reply.

- Central banks have **developed their own data platforms to deal with regulatory data analytics**, but this is often used in conjunction with Excel and document-based handling. This was the answer from over three-quarters of respondents, the majority of which were from developed countries. Only four of the 32 above use purely self-developed regulatory data analytics while the remaining 28 combine it with another standard. Nearly all combined this with Excel and document-based handling and just under half reported a commercial element to their regulatory data analytics workflow.

- Commercial solutions from the market proved popular among respondents, with over half choosing this option. Interestingly, **more central banks use commercial solutions from the market** for regulatory data analytics than they do for regulatory data collection, with only 16 central banks.

BearingPoint.

As of January 2016 Globally Systemically Important Banks (G-SIBS) have been regulated to meet the BCBS 239 Principles, after carrying out self-assessments, in respect of data collection, data aggregation and dissemination capabilities. Do you think it would be useful for central banks to self-assess using an adapted version of these principles?

Seven respondents did not reply.

- Central bankers largely welcome the idea of self-assessment using an adapted version of the BCBS 239 Principles for Globally Systemically Important Banks (G-SIBs). Of the 35 respondents, 80% said such self-assessment would be useful.

- As databases increase in size and complexity, there is naturally a concern that they are policed and governed properly. Furthermore, one central banker noted existing systems were under pressure: "Due to increasingly large and complex data that is now challenging traditional database systems."

- Several central banks drew attention to the internal processes already in place to regulate big data processing. One European respondent from an advanced economy noted: "There is a benefit that can be accrued from this. It is important to note however that our internal audit function already undertakes full audits in this area." However, if the objective of these principles is to strengthen banks' risk data aggregation capabilities and internal risk reporting practices, it might be relevant to make a similar self-assessment in central banks.

BearingPoint.

Results of survey of how central banks view big data and data governance

Summary of key findings (1/2)

- Central banks have an **active interest** in big data. This is manifested in improving processing technology, adapting institutional strategies and increasing staff awareness of the area.

- Central banks typically see big data as **unstructured data** that is sourced **externally**, though this view is not universally held.

- Overwhelmingly, central banks develop their **own data platforms** to handle regulatory data collection, a role that has taken on greater significance since the financial crisis as central banks have expanded their involvement in financial stability.

- Big data is predominantly regarded as **useful for research**, but significant minorities see immediate involvement in policy-making, or scope for this.

- **Lack of support from policy-makers** is seen as the most significant challenge to increase use of big data.

- Central banks **do not in the main have a dedicated budget** for the handling of data (including big data), though many are seeking one.

- A little over 80% of respondents said they **do not have any intra-departmental or divisional bodies** dedicated to big data.

- More broadly, central bankers have **concerns** over the arrangements in place **for managing data** in their institutions. Many are looking to improve data governance.

BearingPoint.

Results of survey of how central banks view big data and data governance

Summary of key findings (2/2)

- Over three-quarters of respondents indicated they had a **shared internal platform**, typically in the form of reporting frameworks and data warehouses.

- Central banks generally **source their own big data sets**, though a significant minority increasingly look elsewhere for these. Overwhelmingly, they process these themselves and there is no indication of a desire for this to change.

- **Monetary policy** is seen as standing to benefit most from big data, though it is expected to have a significant impact on macro-prudential policy as well.

- **Support from the executive-level and policy-makers** divided respondents: 35% saw it as top priority for investment within the central bank, 38% saw it as the lowest.

- Central banks have developed their **own data platforms to deal with regulatory data analytics**, often used in conjunction with other options. Excel remains popular.

- Central bankers broadly **welcome the idea of self-assessment of data management** using an adapted version of the Basel Committee's BCBS 239 principles for supervisory data aggregation.

BearingPoint.