A RESEARCH PAPER PRODUCED BY
FINEXTRA IN ASSOCIATION WITH PRIVITAR

PRIVITAR
ENGINEERING PRIVACY

TACKLING DATA PRIVACY TO UNLOCK THE POWER OF BIG DATA ANALYTICS

Finextra

01
FOREWORD
TURNING DATA INTO VALUE

By Jason du Preez, CEO, Privitar

Today, every business is a data business. Early stage businesses often start to become truly valuable when they have accumulated enough data to create insight, and large, incumbent firms are increasingly recognizing the potential in the data they already hold.

But responsible use of this data is essential. Financial institutions need to know the data they have and how it is being used. To maintain trust, they need to provide clarity and transparency on this to their customers. Without this trust, there is a risk that data quality and supply will be eroded – whereas improved trust and customer relationships lead to better quality data collection and a positive sum game.

While customer trust is generally the primary motivating factor, organizations must also comply with the evolving demands of regulation in data protection. The imminent arrival of the new General Data Protection Regulation (GDPR), designed to harmonize requirements for entities operating across Europe, will bring enforcement of this best practice and severe penalties for non-compliance.

Companies that get ahead of the game stand to gain significantly, generating competitive advantage on the one hand, and avoiding the serious downside of non-compliance with GDPR on the other. This is backed up by the research and interviews on which this paper is based, that clearly show the significant focus on data innovation, alongside awareness that investment in privacy controls can make these endeavors more efficient.

Needless to say this is likely to require transformation programs across people, process and technology. The good news is that smart use of technological controls, with the commensurate rise of the regtech ecosystem, can make this change much more cost-effective and efficient. This approach also results in infrastructures that are easier to manage and scale, driving down long-term total cost of ownership.

Privacy and data protection policies can be supported and enforced with privacy engineering products enabling the creation of useful data that is optimized for specific uses, safely and without privacy risk. This approach, in conjunction with leading data security practices, enables organizations to address the apparent conflict between a desire to exploit data and the necessity of protecting sensitive information.

Privitar is working with leading financial services firms to establish data provisioning pipelines that protect sensitive data in line with clearly defined policies. These ensure transparency and auditability, and empower organizations to broaden access to datasets for R&D, create new data products, share data securely and leverage public cloud for powerful, model-driven applications. In turn these developments also underpin both innovation and compliance.

With financial institutions already seeing returns on investment in modern business intelligence techniques, further investment in data innovation during the next two to three years is likely to be significant. This is confirmed by the expected growth of big data platforms such as Hadoop – predicted to see a 60% CAGR out to 2020.

To make this investment pay in full, protecting data privacy is essential. Emerging technology solutions have a key role to play in helping financial institutions to meet compliance obligations and preserve customer trust, while at the same time unlocking the true value of the data they hold to drive new revenues and sharpen competitive edge.

"Privacy and data protection policies can be supported and enforced with privacy engineering products enabling the creation of useful data that is optimised for specific uses, safely and without privacy risk. This approach, in conjunction with leading data security practices, enables organisations to address the apparent conflict between a desire to exploit data and the necessity of protecting sensitive information."

02 INTRODUCTION

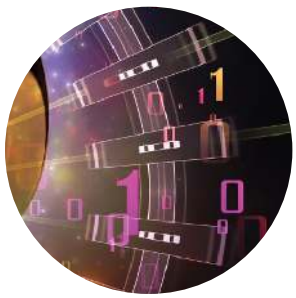
Many financial institutions know that data is their most valuable asset, and are investing in next generation technologies to better leverage that data. This could be to meet obligations in areas such as regulatory compliance, or to go further – capitalize on the data they must capture and analyze anyway to drive secondary commercial goals through sophisticated marketing approaches.

However, many also face challenges when it comes to making the best use of their data assets. Data privacy concerns can limit what banks are able to do with their data – and adhering to privacy rules can slow down the availability of data to underpin innovative new developments. The changing regulatory landscape, with the new GDPR on the horizon, adds a further layer of complexity around what banks can do with data. At the same time, customers' expectations around how their data is used are changing.

Overall, the power of big data analytics is clear, but some fundamental building blocks need to be in place if banks are to unlock that power for commercial benefit.

In this context, this Finextra paper, produced in association with Privitar, sets out to assess how realistic it is for financial institutions to balance successfully the opposing forces of powerful big data analytics and increasingly stringent data privacy obligations.

The paper explores the progress banks are making with big data and in which areas. It examines the challenges most often impeding progress, and identifies the biggest opportunities banks see to do more. It also looks at the value and contribution of technology solutions to enable banks to embed privacy and data protection into their big data approaches.



The paper combines the findings of a detailed online survey carried out by Finextra during late 2016 and early 2017 with the insights of data privacy and protection experts gathered by Finextra through one-to-one interviews.

03
BIG DATA VERSUS DATA PRIVACY?

Achieving the fine balance between exploiting the full potential of big data analytics and complying with data privacy rules is a difficult exercise for financial institutions. Indeed, there is an inherent contradiction at play here, as Tanguy Van Overstraeten, Global Head of Linklaters' Privacy and Data Protection Practice, points out.

"As banks become more digital, they become even more data-rich, and they are dealing with some data that has high value for customers and can be very personal. Beyond that, the banks are developing several new services, usually big data-based, for which they are analyzing large amounts of this information.

> **"With the growing focus on big data and artificial intelligence (AI), we see firms trying to exploit the benefits data can bring, while at the same time remaining compliant with the current framework, and having an eye on the new framework."**

"In effect, their business models are increasingly being driven by customer data, while the framework being conceived by regulators is shaped by a drive to reduce the use of this data. I can see a real contradiction between what the regulators want – data minimization and additional layers of protection – and the new services banks and many other sectors are developing which will benefit not only the banks but also consumers."

In Europe, with the recent adoption of the GDPR and its application as of 25 May 2018, this contradiction is even more exacerbated, says Van Overstraeten, it will need to be tackled if the vision of a digital Europe is to be realized. "Data privacy must be regulated, of course, but in a way that does not hamper the development of business and innovation," he adds.

> "It is not a case of big data 'or' data protection, or big data 'versus' data protection. That would be the wrong conversation. Privacy is not an end in itself: it is an enabling right. Embedding privacy and data protection into big data analytics enables not only societal benefits such as dignity, personality and community, but also organizational benefits like creativity, innovation and trust. In short, it enables big data to do all the good things it can do."
>
> ELIZABETH DENHAM, INFORMATION COMMISSIONER

David Smith, former Deputy Commissioner at the UK Information Commissioner's Office (ICO), and Special Adviser to Allen & Overy, agrees. "With the growing focus on big data and artificial intelligence (AI), we see firms trying to exploit the benefits data can bring, while at the same time remaining compliant with the current framework, and having an eye on the new framework."

However, the challenge is not confined to GDPR and Europe. As Van Overstraeten says, some 120 countries around the world now have data protection rules. "There are three broad categories. Those which mirror Europe – which would include some countries in North Africa, and South Africa, which has even pre-empted GDPR with a law inspired by it now progressively entering into force. Then there are the countries with some rules but which are quite general, like India and China. The third category includes countries with even more strict rules, such as Korea, where the approach can be very stringent. The US is a kind of patchwork, with a sectorial approach and several State laws with diverse levels of requirements."

To add another layer of complexity, the reach of EU data protection law has been shown to extend to businesses offering services in Europe, even if they are not headquartered in Europe, by the outcome of legal cases such as that of Google Spain. As Smith points out: "The new regulation reinforces that territorial reach, but the courts have gone a long way to establishing this already, and the regulation has simply followed."

Smith: Firms balancing need to exploit data benefits with data privacy compliance

The data privacy challenge for banks is not confined to their commercial activities. In the compliance space, too, banks can be caught between a rock and a hard place. "Anti-money laundering regulation requires banks to put a lot of emphasis on identifying people involved in deals, but this still has to be proportionate and reasonable," says Smith. "When investigations start to extend to people who are not directly involved, but are family members of those who are, this can be viewed as intrusion."

Cameron Craig, Deputy General Counsel, Data Privacy and Digital, Group Head of Data Privacy, Group Legal, HSBC, agrees "there is a tension between our need to use data for the purposes of detecting financial crime and data privacy requirements in some countries". "Fighting the bad guys, complying with AMLD4 and FATCA, all drive towards the central processing of data and the building of a global picture of activities, so that we can do that more effectively," he says.

"On the other side, data privacy obligations, data localization requirements and bank secrecy legislation in some countries make it difficult to easily share data. This is a policy area in which HSBC and others are working to try and achieve greater co-operation between governments and regulators to find ways to allow banks to more effectively share data both within a corporate group and also between corporate groups and between the private sector and law enforcement and government," Craig adds.

The good news is that this contradiction is well recognized and thought is being put into how to resolve it. The UK ICO recently published an updated paper on *Big data, artificial intelligence, machine learning and data protection*, and in the paper's foreword Information Commissioner Elizabeth Denham acknowledges the challenge: "It's clear that the use of big data has implications for privacy, data protection and the associated rights of individuals – rights that will be strengthened when the GDPR is implemented. Under the GDPR, stricter rules will apply to the collection and use of personal data. In addition to being transparent, organisations will need to be more accountable for what they do with personal data. This is no different for big data, AI and machine learning."

However, Denham also emphasises that "implications are not barriers". "It is not a case of big data 'or' data protection, or big data 'versus' data protection. That would be the wrong conversation. Privacy is not an end in itself: it is an enabling right. Embedding privacy and data protection into big data analytics enables not only societal benefits such as dignity, personality and community, but also organisational benefits like creativity, innovation and trust. In short, it enables big data to do all the good things it can do."

Craig: There's a tension between our need to use data for the purposes of detecting financial crime and data privacy requirements in some countries

04

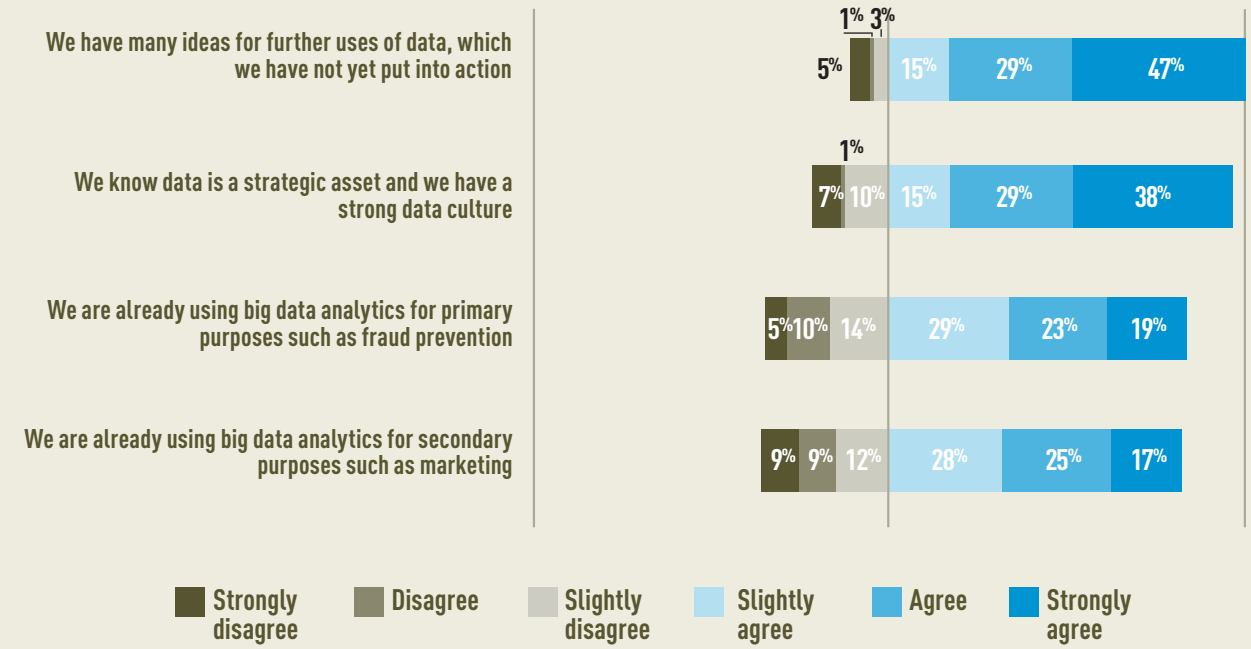
SO HOW DO FINANCIAL INSTITUTIONS CURRENTLY RATE THEIR BIG DATA ANALYTICS PROGRESS?

The results of the online Finextra survey – completed by 150 respondents from 84 institutions across 22 countries – paint a pretty positive picture upfront of the status of firms' big data approaches, and their thinking about the value and potential of big data.

As Chart 1 shows, a substantial 67% agree or strongly agree with the statement "we know data is a strategic data asset and we have a strong data culture". Further analysis reveals that – encouragingly – having a strong data culture isn't just about size of institution or department budget. Those with an annual budget of up to €1M were only slightly less likely to claim a strong culture than those with budgets over €10M.

CHART 1

PLEASE INDICATE HOW WELL THE FOLLOWING STATEMENTS REFLECT THE CURRENT STATUS OF, AND THINKING ABOUT, BIG DATA ANALYTICS WITHIN YOUR FIRM.



Meanwhile, 76% agree or strongly agree that "we have many ideas for further uses of data which we have not yet put into action". At the very least this suggests a great deal of creative thinking is under way – and it could also imply a strong investment pipeline in this area.

The findings around firms' current progress in using big data for primary (mandatory, such as fraud prevention) and secondary (optional, such as marketing) purposes are more muted, but still positive, with 71% agreeing that they are already using data analytics for the former and 70% for the latter.

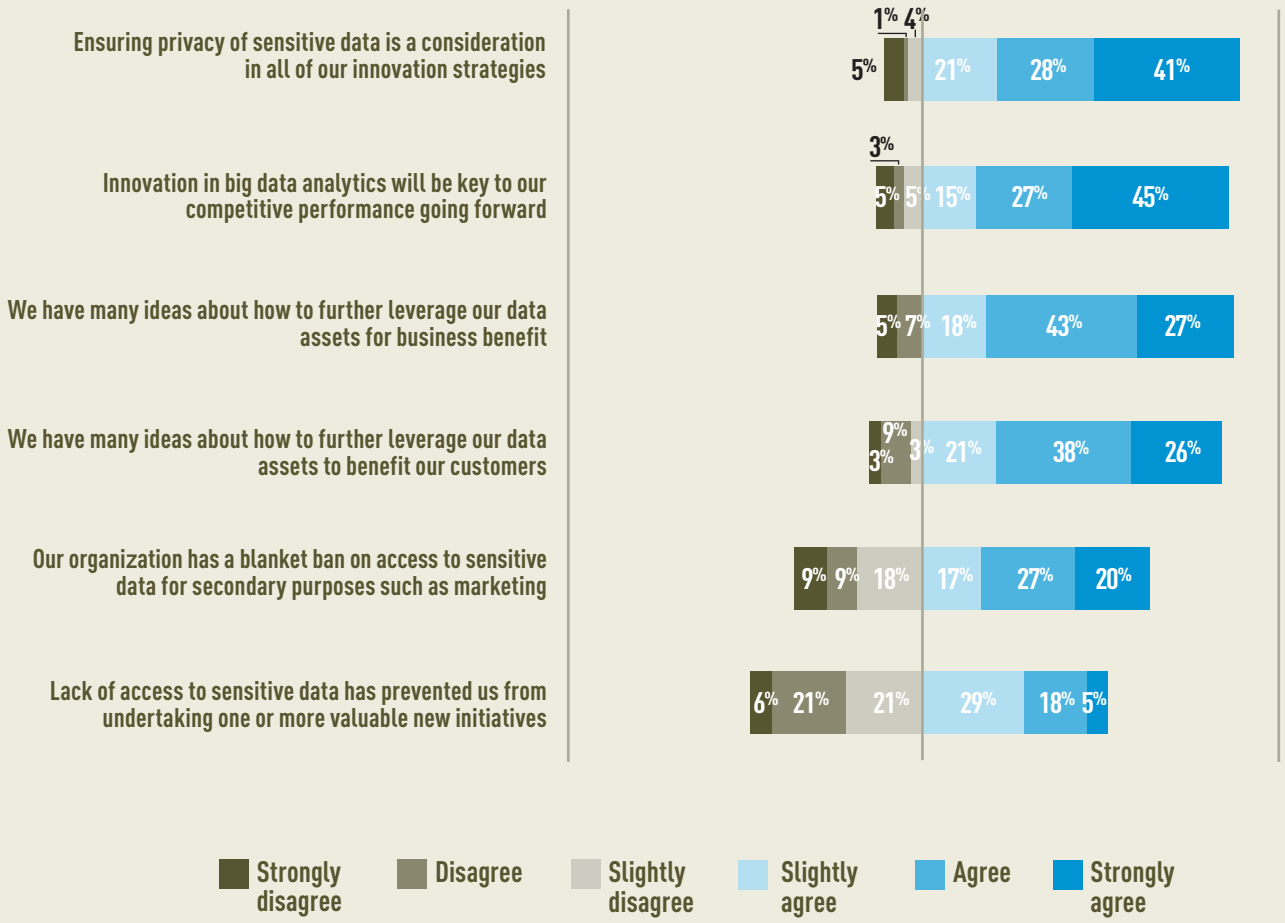
Drilling down more deeply shows that UK and North American respondents are more likely to already be using big data analytics for secondary purposes (50% and 47% agree or strongly agree), compared to 35% in Western Europe and 20% in APAC.

All that said, overall, less than half of all respondents agree or strongly agree they are using big data analytics either for primary or secondary purposes – with the strongest disagreement (9%) being registered for secondary purpose applications. This could indicate that a lot of projects are still in the planning or early deployment phases. It could also reflect the challenges firms face in striking the balance between what they want to do with big data, and what they must do to protect data privacy.

Chart 2 offers further insights on this question of how easy it is for firms to utilize data for secondary purposes. It shows that a significant 47% agree or strongly agree that they simply are not allowed to use sensitive data in this way, with regional drill-down showing that this agreement ranges from 35% in North America to 55% in APAC. Importantly, 52% of respondents agree in some way that lack of access to sensitive data has prevented them from undertaking one or more valuable initiatives. The degree to which this was agreed with varied significantly from region to region, with only 5% of UK respondents agreeing or strongly agreeing, compared to 40% in APAC and 25% in Western Europe.

CHART 2

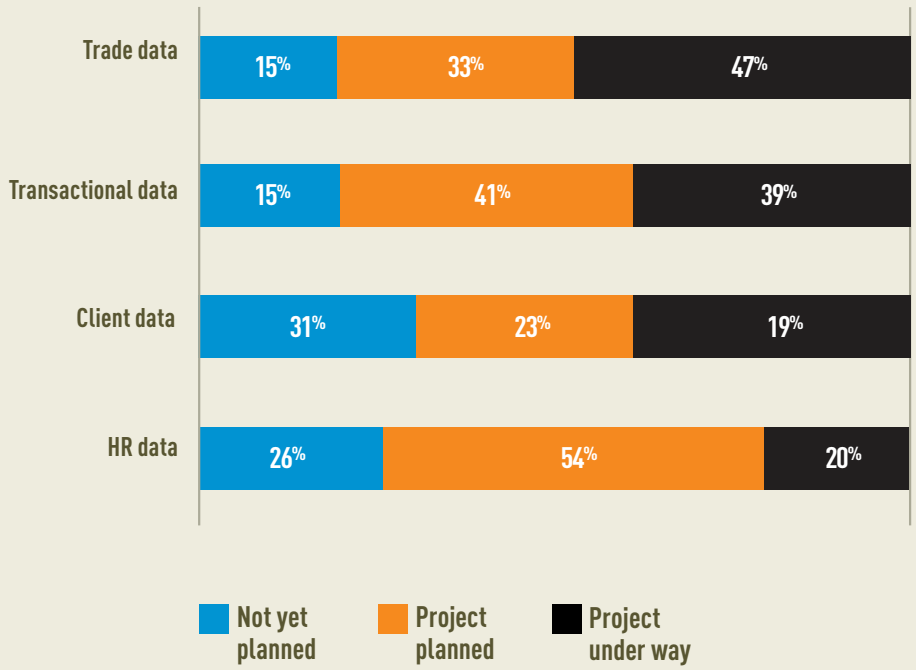
PLEASE INDICATE YOUR LEVEL OF AGREEMENT WITH THE FOLLOWING STATEMENTS ABOUT INNOVATION AROUND DATA WITHIN YOUR FIRM.



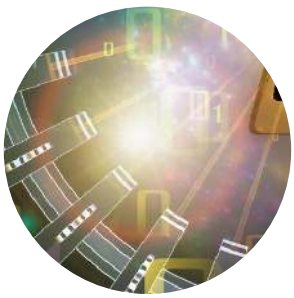
The relative difficulty of carrying out big data analytics on personal data could also be reflected in the findings shown in Chart 3, which presents respondents' input on the types of data they are currently using for big data analytics. Though the variation in use of the different data types is not very marked, it is worth noting that client data is the data type for which the biggest percentage of respondents indicated they have no projects yet planned (31%). HR data is the type for which the smallest proportion indicated they have projects already under way (20%).

CHART 3

PLEASE INDICATE WHETHER YOU HAVE CURRENT OR PLANNED PROJECTS TO UTILIZE THE BELOW TYPES OF DATA FOR BIG DATA ANALYTICS



"Client data is the data type for which the biggest percentage of respondents indicated they have no projects yet planned (31%). HR data is the type for which the smallest proportion indicated they have projects already under way (20%)."



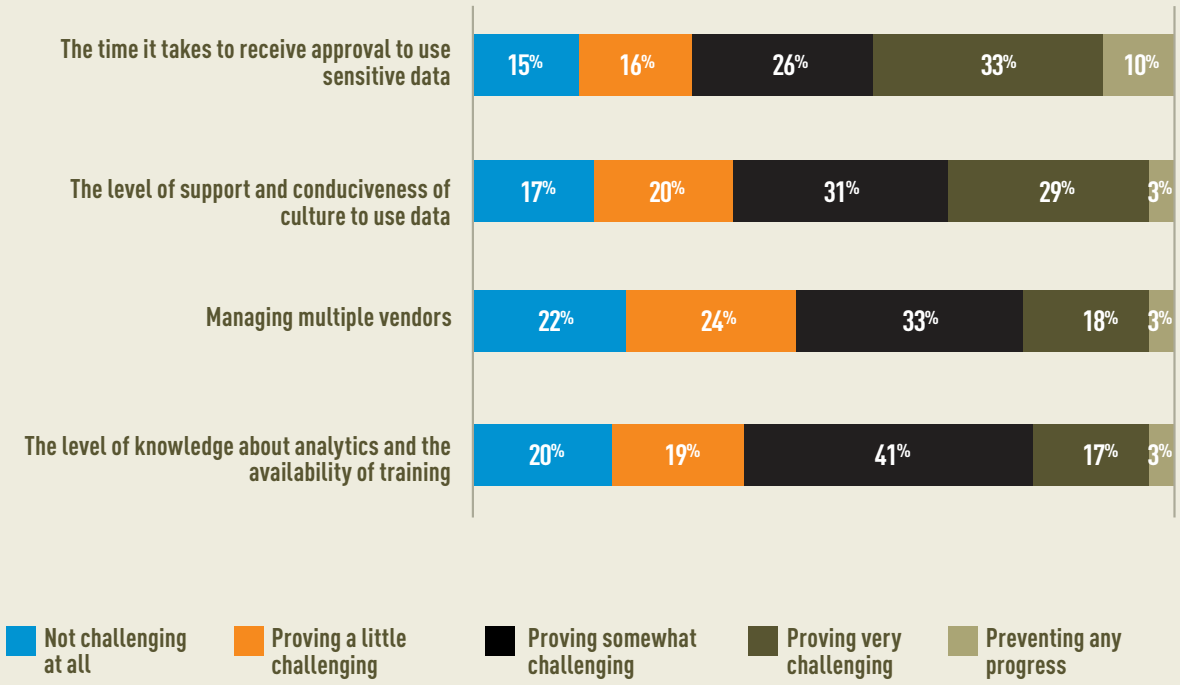
05 WHAT ARE THE CHALLENGES HOLDING UP FIRMS' BIG DATA PROGRESS?

INTERNAL BARRIERS

Asked to identify the biggest internal barriers to achieving their big data vision, respondents to the online survey indicated that "the time it takes to receive approval to use sensitive data" is the most problematic issue (Chart 4). This was described as very challenging by 33% of respondents, and for 10% it is preventing any progress at all.

CHART 4

PLEASE INDICATE HOW MUCH THE FOLLOWING INTERNAL FACTORS ARE CREATING CHALLENGES FOR YOUR FIRM IN REALIZING ITS BIG DATA ANALYTICS VISION.



While regulation – current and future – "doesn't impose a timeframe" on the processes banks go through to establish whether data can be used, as Smith says, "businesses can take a long time to assess the data privacy implications of activities, as these things require thinking through". "It's important that financial institutions always keep in mind the privacy impact of what they are doing, and take a balanced approach," Smith continues.

"The idea is not to exploit every new opportunity without thinking. Everything should go through a process, and with every new development, firms should ask themselves, can we deploy technology to help protect privacy, perhaps by pseudonymizing or anonymizing the data? And it's not just the results that matter. Banks need to be able to show the process they have been through if they are challenged by the regulator. Even if the regulator takes a different view of the outcome, if you have followed a rigorous process this will help in mitigation," he adds.

Good practice already, privacy impact assessments become mandatory under GDPR, which also intensifies the focus on being able to demonstrate what has been done. This may not – in the short term at least – speed up decision-making, but it is nonetheless a "good thing", says Craig, "because it's making banks focus on what everyone is referring to as the 'data mapping exercise'. This involves really making sure that we understand what data we hold, what we do with it, what the reason for doing that is, and then if consent is required, making sure we get clear consent, and notify individuals appropriately. It's pretty straightforward what we have to do – but doing it in practice, in a way we can demonstrate, is the challenge."

The stakes under GDPR are so high that cutting corners would be ill-advised: where previously banks may have been more willing to take a risk-based approach to data privacy compliance on the basis that the potential sanctions were not that great compared to the penalties applicable under other financial services regulation, in the new landscape, the threat of fines to the tune of 4% of global group turnover clearly demand the application of the most rigorous risk assessment procedures.

Nonetheless, the increased rigor and demonstrability required by GDPR can also be considered as positive developments, insists Van Overstraeten.

Van Overstraeten: Rigor demanded by GDPR is a positive development

"The second most troublesome internal barrier is "the level of supportiveness and conduciveness of culture to use data", which is ranked as a challenge to some degree by 83% of respondents – with North Americans in particular highlighting this barrier."

"The opportunities are numerous," he says, "and one is certainly that for the implementation of the GDPR stakeholders must find ways to streamline data flows, to be able to identify and organize them on a more central basis, because it will become very difficult to address the new requirements on a local basis. So, banks must take a top-down approach, which means in terms of management and organization this should have a positive impact and lead to more efficiencies."

Getting permission to use sensitive data is not the only impediment to speedy implementation of data analytics, as Mark Ainsworth, Head of Data Insights, Schroders Investment Management, explains. "Once we find out from an investor the question they want to answer, we have to find the data source, obtain it, make it technically stable, refine it to deliver the insights we need… this can easily take months, and investors usually want it in weeks," he says. Schroders now has a number of ready-to-use data assets which can speed this process, but, sensitive data or not, "it takes a certain amount of time" to answer the questions investors raise, Ainsworth emphasizes.

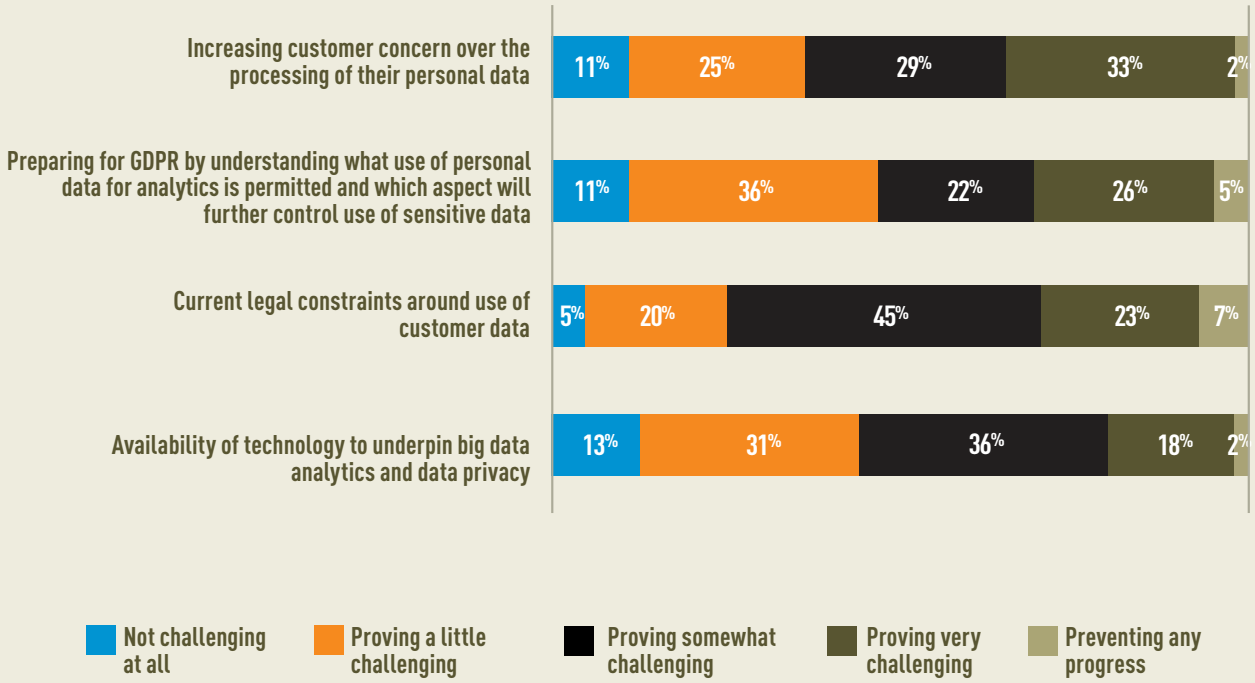
Behind the time challenge, Chart 4 shows that the second most troublesome aspect when it comes to internal barriers is "the level of supportiveness and conduciveness of culture to use data", which is ranked as a challenge to some degree by 83% of respondents. Interestingly, North Americans, in particular, highlight this barrier, with 41% indicating they find support and cultural conduciveness very challenging, and 6% saying it prevents any progress at all.

There are cultural and communications challenges to be negotiated, agrees Ainsworth. "Our data scientists need to be able to articulate the insights in the data, so investors don't have to struggle to understand it," he says. "The insights need to be relevant, true, useful and actionable. They also need to be novel. This is not an intrinsic quality of data. This means we need to work in partnership with the investors. We are not an ivory tower. We can't tell what is insightful without knowing what the analysts already know – so we can move their understanding forward. The other challenge is that every team is different. This internal diversity means that every team needs us to interact with them in a different way."

Ainsworth: Cultural and communications challenges must be overcome

CHART 5

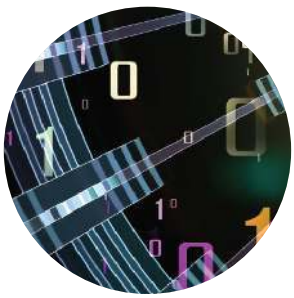
PLEASE INDICATE HOW MUCH THE FOLLOWING EXTERNAL FACTORS ARE CREATING CHALLENGES FOR YOUR FIRM IN REALIZING ITS BIG DATA ANALYTICS VISION.



EXTERNAL BARRIERS

When it comes to external barriers to the realization of firms' big data analytics strategies, as Chart 5 shows, the survey respondents indicate "current legal constraints around the use of customer data" is the biggest. The smallest proportion say it is not challenging at all (5%), the biggest proportion say it is preventing any progress (7%), and a significant 68% describe it as very or somewhat challenging.

This is clearly consistent with the findings around internal barriers, the time it takes to get approval for use of sensitive data for data analytics projects being directly impacted by the current legal constraints around the use of customer data.



The second biggest barrier identified by the survey respondents is "increasing customer concern over the processing of their personal data", with 89% saying this is challenging to some degree, and 33% describing it as very challenging.

Growing customer awareness of the value of their data is indeed a critical factor in play here. As Ainsworth says, so much of the value of big data comes from the ability for companies to personalize products and services based on what the analytics tell them about customer requirements, but customers know this as well as banks do. "My personal view is that I would much rather there was a more explicit open value exchange," he says. "I get a good free email product, which otherwise I might pay for, in exchange for my data being used. I get a lower tariff for phone calls in exchange for being marketed to. I benefit from money off at the supermarket in return for allowing my data to be processed. There should be an open value exchange, into which customers freely enter."

> **"Once we find out from an investor the question they want to answer, we have to find the data source, obtain it, make it technically stable, refine it to deliver the insights we need... this can easily take months, and investors usually want it in weeks."**
>
> MARK AINSWORTH, HEAD OF DATA INSIGHTS, SCHRODERS INVESTMENT MANAGEMENT

To make life more difficult for banks, customers' expectations of the extent to which they should protect their data – and the degree to which they should use it responsibly – exceed those they apply to other businesses, suggests Mark Mullen, CEO, Atom Bank. "Money is a pretty emotional construct, and a bank therefore has a fundamental responsibility which is, don't lose your customers' money. In today's society, you could equally attach the same meaning to the words, don't lose your customers' data," he says. "Therefore it would be a mistake to believe that the standards customers expect from other types of site or other types of internet environment or app-based environment would be acceptable for banks."

On the contrary, there needs to be "state-level security protocols that all banks must comply with", Mullen adds. "These also should enshrine protections reflecting the fact that data in and of itself has a value. In the UK, we have the Financial Services Compensation Scheme, so that if someone loses money, there's recourse. But there's no recourse for customers who lose their data, and yet – everyone will tell you – that data has a strategic value. If that's the case, then we must monetize it. What does the customer get if something happens to their data?"

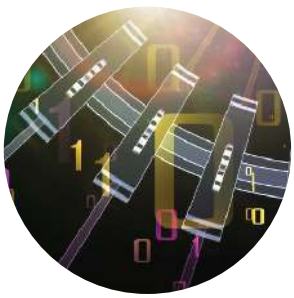
> "We can use evolving technology in a way that helps both the bank and our customers, and we can only do that if we are data privacy compliant, if we're clear on what customers are consenting to, if we're clear that we are very transparent with our customers as to how we use their data, and we're clear that what we're doing with data is fair, lawful and consistent with our conduct responsibilities."
>
> CAMERON CRAIG, DEPUTY GENERAL COUNSEL, DATA PRIVACY AND DIGITAL, GROUP HEAD OF DATA PRIVACY, GROUP LEGAL, HSBC

Mullen further highlights the importance of customer consent. "Does the customer absolutely understand what you are doing with their data? Has the customer given you explicit permission to use their data in the way you are using it? Can you prove it? One of the things that I am concerned about is how we can actively demonstrate that people understand what they are signing up to. I remain very skeptical about the conscious acceptance of T&Cs. I think the industry still has an awful lot to do to simplify the language it's using."

Van Overstraeten raises a similar point. "We all see lots of these privacy policies around and to be frank hardly anyone reads them. The EU Commission is considering a system of icons, which children could learn about at school, to enable consumers to immediately understand what will be done with their data. This of course must be the work of public bodies, not private companies."

Smith agrees "there is an educational aspect to this", adding: "We will certainly see a move away from the traditional T&Cs approach, towards getting permission and consent at the right time, similar to the way in which you might be asked whether you want to permit the use of location data on your phone," he says.

He too highlights the growing importance of consent, which is strengthened under GDPR. "If you are tracking data to build up a picture of people's lives to offer them better, targeted services, they are likely to welcome it, assuming they have consented. It has to be a real, clean choice for customers though. Banks can certainly offer their customers many great new products and services, but they must get them to sign up first, in simple terms, that customers can understand," Smith says.



Avoiding alienating customers with data use is essential, Smith adds. "The greatest risk is if you upset people. Don't take customers by surprise. If they think you're doing something remiss, this will trigger complaints – and that is what creates regulator interest."

Techniques such as anonymization and pseudonymization come into play here, he continues. "You may well be able to do the analysis you need using anonymized/pseudonymized data, because you don't always need to know real world identities to generate useful insights. Using pseudonymization also reduces the risk of data being wrongfully accessed, thus minimizing data breaches – mitigating the huge risk of hacking when substantial amounts of data are collected together."

Once again though, this challenge – customer perceptions about data use – can also be turned into an opportunity, according to Craig. "We can use evolving technology in a way that helps both the bank and our customers, and we can only do that if we are data privacy compliant, if we're clear on what customers are consenting to, if we're clear that we are very transparent with our customers as to how we use their data, including providing them with choice over how we use their data where appropriate or required, and we're clear that what we're doing with data is fair, lawful and consistent with our conduct responsibilities," he says.

"In this context, the introduction of privacy impact assessments and privacy by design is helpful. For example, we might anonymize or pseudonymize data to use it, and the rigorous approach of further embedding privacy into the project lifecycle can actually enable us to do things with data that we'd otherwise not be able to do for both our benefit – and the customer's benefit."

> **"One of the challenges is that when you talk about data protection, you enter a field that for many years not so many people were dealing with, and therefore the level of knowledge is usually relatively low – although it is improving – and you are faced by a very sophisticated set of rules: 99 articles, representing about 100 pages of legislation, and 173 Recitals to help the interpretation of the rules."**
>
> TANGUY VAN OVERSTRAETEN, GLOBAL HEAD OF LINKLATERS' PRIVACY AND DATA PROTECTION PRACTICE

THE ELEPHANT IN THE ROOM: GDPR

"Preparing for GDPR" is also identified as a challenge by 89% of respondents, though a smaller proportion (48%) describe it as very or somewhat challenging. For impacted institutions – as previously mentioned, not just those domiciled in impacted European markets – the size of the task of preparing for GDPR should not be underestimated. All the observers interviewed for this paper confirm how high a priority GDPR is being afforded by firms.

Van Overstraeten describes the GDPR as "the most critical" driver of change in the data protection landscape. "It's a huge piece of work," he says. "We have started to work with a number of clients on this. It takes a lot of time, and therefore businesses need to move really quickly. Every time I am in a meeting and I talk about data protection, even CEOs who have no clue about legal aspects refer to GDPR, anywhere in the world. This is really a boardroom issue."

Smith agrees, adding that "businesses at the top of their game should have a programme in place – which takes resources, time and effort, internally and externally". "Those that have traditionally been ahead of the game, and doing a good job under the current law, will find it easier," he continues. "We see a mixed picture at the moment, with some firms well-positioned, and others dragging their feet, hoping it won't happen: but it will."

That said, the size of the project is such that "it is virtually impossible to do everything by May 2018", Smith adds, "so firms' programmes must be built around the main priorities and an assessment of risk."

The complexity of GDPR is also an important consideration, Van Overstraeten continues. "One of the challenges is that when you talk about data protection, you enter a field that for many years not so many people were dealing with, and therefore the level of knowledge is usually relatively low - although it is improving - and you are faced by a very sophisticated set of rules: 99 articles, representing about 100 pages of legislation, and 173 Recitals to help the interpretation of the rules. Each time I read back one of these provisions I see more complexity, though I have already spent four years dealing with the drafts which became the final GDPR. So one big challenge is that GDPR is now in a sphere of real experts and it's something that will not be accessible to all."

> "In the past, the difficulty we have faced is having data in multiple systems, organized in different departments, set in silos. Organizations had the data, but they were not able to exploit it in an agile way. The cost was prohibitive. New technology such as data lakes allows us to group data from multiple systems into a single repository and this is helping institutions to be much more responsive and agile. The data is then being extracted to meet specific needs, irrespective of the fact that it was being handled by different systems. This means we can be much more agile."
>
> MATHIEU MAURIER, GLOBAL HEAD OF SALES AND RELATIONSHIP MANAGEMENT, SOCIETE GENERALE SECURITIES SERVICES

That said, the very size and complexity of the GDPR compliance challenge brings benefits, he suggests. "GDPR is an eye opener for many businesses," he says. "Before this, the enforcement around data protection breaches was relatively low. The environment is changing rapidly which means that businesses can push this high on their agendas and can approach it in a structured manner and with a positive mindset. An increasing number of companies explore the marketing aspects of being compliant and seize that as an opportunity to improve their image towards their customers and their personnel. In the current legal regime, my overall impression was that many businesses were simply complying with the formalities as an administrative process and then forgetting about data protection. With the GDPR it's a continuous process. They will have to update everything they do on a regular basis to make sure they continue to be compliant - which is good, as long as it is not too burdensome."

Balance is key, Craig agrees, also emphasizing the alignment between what GDPR demands, and what banks should be doing anyway to ensure fair treatment of their customers. "The balance we need to achieve is not just about solely complying with data protection law – because that's not the culture within banks now," he says. "We look as much to conduct outcomes, which very much plays into decisions about how we use data. We want to use customer data lawfully and we want to use customer data in a way in which our customers would expect us to use it.

"Maintaining that balance means we need to do a privacy impact assessment, to make sure that when we are using data to introduce a new product or service or systems we have that 'stop and think' moment. What are the privacy implications of this? Is it lawful, and would the customer expect us to do this? Will it lead to a fair outcome for the customer?

"Steps that can be taken to alter the balance would include asking whether we could achieve the same goal using less customer data, allowing fewer people to access the data, or masking or anonymizing the data? This is embedding privacy by design concepts, to try and tip the balance more towards protecting customer data privacy."

TECHNOLOGY: A SILVER BULLET?

The least challenging external factor, according to the online survey respondents, is "the availability of technology to underpin big data analytics and data privacy". As Chart 5 shows, for 13% this is not challenging at all – though 67% still said it is a little or somewhat challenging.

CHART 6

PLEASE INDICATE WHICH OF THE FOLLOWING TECHNOLOGY SOLUTIONS YOU HAVE IMPLEMENTED IN YOUR DATA ENVIRONMENT, AND WHICH ARE STRATEGIC.

(I.E. THEY ARE NOT EXPERIMENTAL/ONE-OFF BUT RATHER PLATFORM INVESTMENTS YOU PLAN TO BUILD ON FOR THE FUTURE).

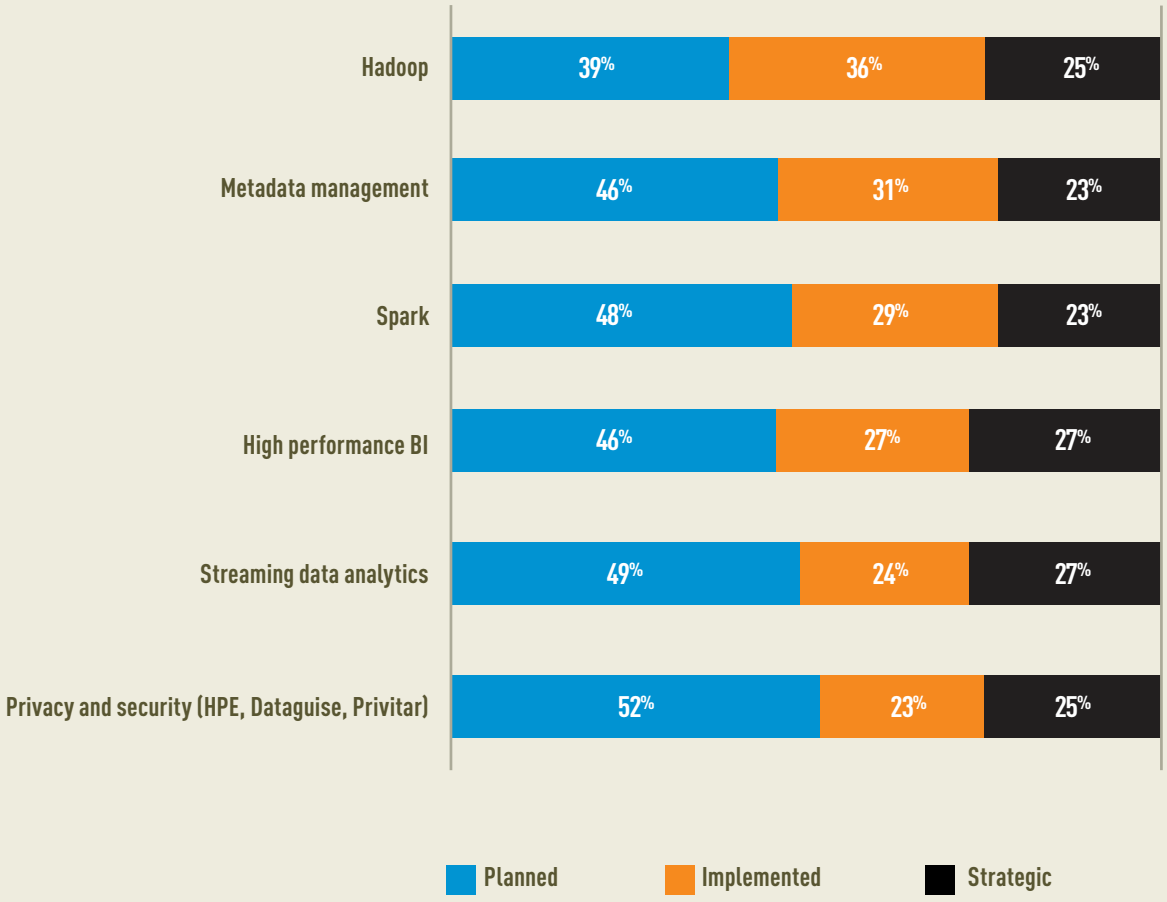


Chart 6 presents further insights from the survey into the current state of big data and data privacy technology implementation within the respondents' firms. As can be seen, Hadoop is the most implemented of the technologies featured, with 61% saying they have put this in place and 25% saying it is a strategic platform. High performance BI and streaming data analytics come out as the most strategic – just – at 27%, while privacy and security technology solutions rank as the most planned but least implemented: just under half (48%) have implemented specific privacy and security solutions in their data environments.

Not surprisingly, there is a correlation with department budget size and implementation rates of these technologies, and the implementation rate for privacy and security solutions rises to 64% when looking at those with respondents with budgets of more than €5M.

Implementing big data platforms like Hadoop is unavoidable if true big data analytics are required, suggests Ainsworth. "Not all the data sets we use are big data, but some of them are, and for this we have needed to implement solutions that did not exist in the firm before our team was set up," he says. "Traditional databases are just not up to the job of handling large volumes of digital dust trail data, and we are also doing text analytics on sources like annual reports, SEC findings and earnings calls, so to support this we need cloud database technologies able to hold and process the data," Ainsworth adds.

Big data technologies are certainly helping banks to circumvent traditional data challenges, concurs Mathieu Maurier, Global Head of Sales and Relationship Management at Societe Generale Securities Services. "In the past, the difficulty we have faced is having data in multiple systems, organised in different departments, set in silos. Organizations had the data, but they were not able to exploit it in an agile way. The cost was prohibitive. New technology such as data lakes allows us to group data from multiple systems into a single repository and this is helping institutions to be much more responsive and agile. The data is then being extracted to meet specific needs, irrespective of the fact that it was being handled by different systems. This means we can be much more agile," he says.

"Hadoop is the most implemented of big data technologies, with 61% saying they have put this in place and 25% saying it is a strategic platform. High performance BI and streaming data analytics come out as the most strategic – just – at 27%, while privacy and security technology solutions rank as the most planned but least implemented: just under half (48%) have implemented specific privacy and security solutions in their data environments."

Technology of various kinds plays an "incredibly important role" in enabling firms to both perform big data analytics and ensure data privacy, agrees Craig. "To be able to manage data mapping, across the massive number of customers we have, to control updates to that, to allow customers to change their preferences, to give real transparency to customers and to hold data securely, making sure it is only accessed to be used for the purposes which we have agreed it should be, technology is absolutely crucial," he says. "There is no way we would be able to achieve this just by using Excel spreadsheets, so even aside from the clever big data activities, to enable us to do the basics, the proper use of technology is absolutely key."

Technology also has a key role to play in enabling the increasingly important concept of privacy by design, as Dominic Venturo, Chief Innovation Officer, US Bank, told Finextra. "The principle of privacy by design rests on only acquiring the information needed to complete a task, thus vastly reducing the amount of data being stored, and at the same time protecting customer data through anonymization or tokenization," he says.

Van Overstraeten too highlights the value of technologies which offer ways to anonymize data – "to code data, transforming it for statistical purposes so it comes out of scope of the personal data protection rules".

Such solutions do offer options for firms to sidestep some of the more strict and stringent rules, he says – though there is are no silver bullet solution because "under GDPR there is not so much difference between pseudonymized data and personal date per se. Only real anonymization allows to escape the application of the privacy rules".

In addition, firms must consider certain risks when adopting such data masking solutions, as Gartner analyst Marc-Antoine Meunier pointed out in a recent paper entitled, How Data Masking is Evolving to Protect Data from Insiders and Outsiders. Data breaches continue to harm organizations and highlight the need for effective data protection, and in that context security and risk management leaders should ensure data masking is part of their

"The principle of privacy by design rests on only acquiring the information needed to complete a task, thus vastly reducing the amount of data being stored, and at the same time protecting customer data through anonymization or tokenization."

DOMINIC VENTURO, CHIEF INNOVATION OFFICER, US BANK

portfolio of solutions, Meunier says, but he also points out that real-time, dynamic masking of production data is still maturing, and requires careful performance and application architecture considerations. Data masking also carries re-identification risks and needs to be implemented carefully, he adds.

Gartner's recommendations are therefore to:

– Adopt more than one data masking technology (static data masking, dynamic data masking, data redaction, format-preserving encryption and tokenization et cetera). Together, these solutions would cover a broader spectrum of use cases and software lifecycle phases.
– Evaluate solutions that support the portfolio of platforms that process sensitive data within an organization, including relational database management systems, mainframes, big data, unstructured files et cetera. Pay particular attention to big data platform support, and investigate vendor roadmaps for unavailable capabilities.
– Evaluate re-identification risks as part of data masking projects, and consider vendors that offer tools to establish the re-identification risks.

In its recent paper on big data and data privacy, the UK ICO also highlights the potential of anonymization, which it says "can be a successful tool that takes processing out of the data protection sphere, and mitigates the risk of loss of personal data". However, the ICO too points out that: "Organizations using anonymized data need to be able to show they have robustly assessed the risk of re-identification, and have adopted solutions proportionate to the risk. This may involve a range of technical measures, such as data masking, pseudonymization and aggregation, as well as legal and organizational safeguards."

Often associated with the implementation of techniques to anonymize or pseudonymize personal data is the concept of privacy by design, the ICO continues, adding: "The basis of the privacy by design approach is that if a privacy risk with a particular project is identified, this can be an opportunity to find creative technical solutions that can deliver the real benefits of the project while protecting privacy. Implementing privacy by design solutions can be mutually beneficial for individuals, big data organizations and society."

The ICO also emphasizes the need to complement anonymization techniques with a range of other technical and organizational measures, including security measures to prevent data misuse, such as access controls, audit logs and encryption; data minimization measures, to ensure that only the personal data that is needed for a particular analysis or transaction (such as validating a customer) is processed at each stage; and purpose limitation and data segregation measures which ensure that, for example, personal data is kept separately from data used for processing intended to detect general trends and correlations.

Overall, the ICO acknowledges that "more work is needed on privacy-enhancing technologies", but confirms that "the concept of privacy by design is key in identifying the privacy requirements early in the big data analytics value chain, and in subsequently implementing the necessary technical and organizational measures".

Crucially, the concept of privacy by design has now been included in GDPR, the ICO points out, meaning it becomes a legal requirement, and that data controllers must take "appropriate technical and organizational measures for ensuring that, by default, only personal data which are necessary for each specific purpose of the processing are processed".

For Mullen, this enshrines in regulation what should be good business practice. "When it comes to data, don't save it if you don't need it," he says. "Banks should only store what they need. I have yet to meet an organization that puts its hand up to say, we gather all your data so that we can manipulate what we sell to you and ensure that we maximize profitability for shareholders, yet if you think about the trajectory of the last 20 years, time and time again there have been incidences of banks using customer data to try and pump sales. There's a huge temptation to use customer data 'intelligently', that is solely with a view to maximize revenues."

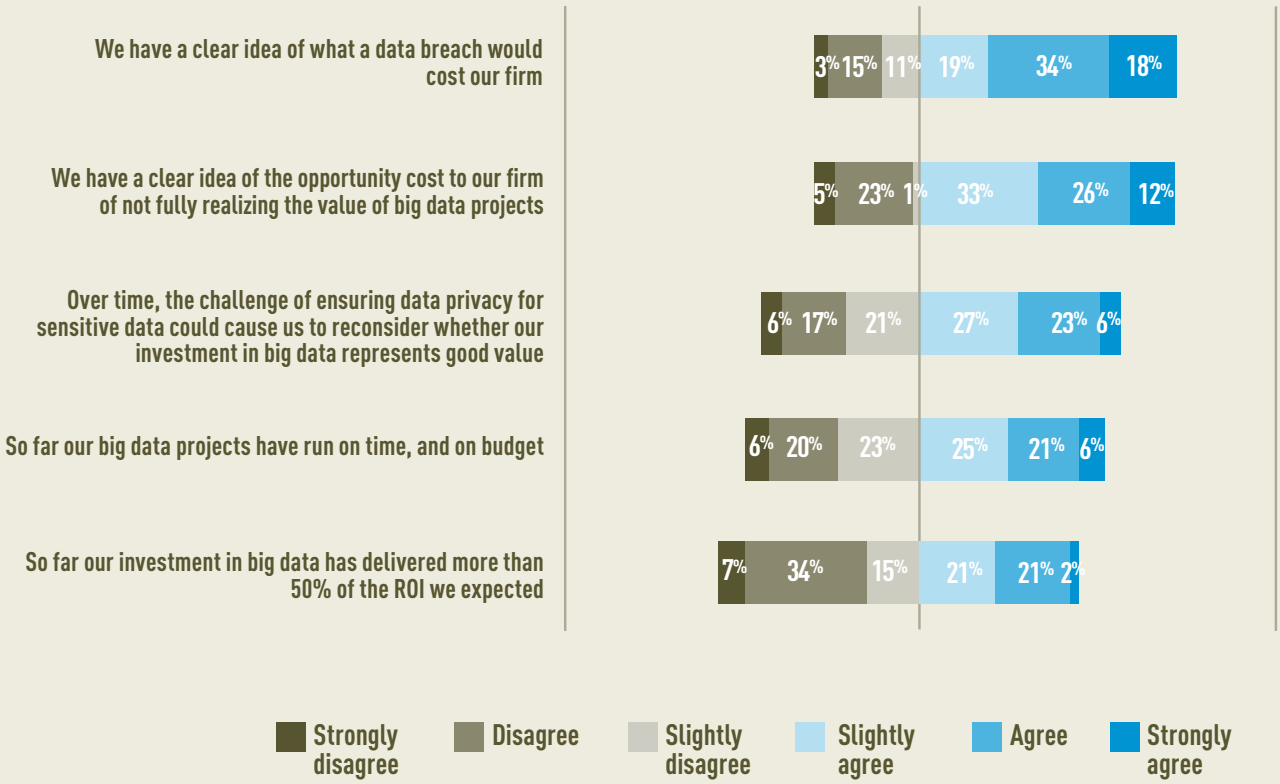
Ultimately though, as Craig points out, privacy by design and technologies that enable the anonymization and pseudonymization of data enable banks to strike that all-important balance. "This allows us to do the profiling and analytics that are not possible with raw data – meaning we can do more advanced profiling, for the bank's benefit but also for our customers' benefit."

06 WHAT IS AT STAKE IF PRIVACY CHALLENGES ARE NOT ADDRESSED?

Chart 7 presents the findings of the survey related to firms' understanding of the value of big data analytics to their businesses – as well as the cost of data privacy-related mistakes. They show that firms have a strong understanding of what a data breach would cost them: 52% agreed or strongly agreed with this statement, rising to 73% for UK respondents.

CHART 7

PLEASE INDICATE YOUR LEVEL OF AGREEMENT WITH THE FOLLOWING STATEMENTS ABOUT MEASURING THE VALUE OF BIG DATA ANALYTICS TO YOUR FIRM.



A smaller percentage – 38% – have the same strong sense of the opportunity cost to their firms of failing to realize the full value of big data analytics projects, though more than half agree in some way that over time, the challenges of ensuring data privacy for sensitive data could cause them to reconsider whether their investment in big data represents good value.

The only statement with which fewer than half of respondents agreed is "so far, our investment in big data has delivered more than 50% of the ROI we expected": in other words, under 50% of respondents have so far reaped more than half of the returns they anticipated. Drilling down into the regional correlations reveals that only 12% of North Americans agreed or strongly agreed with this statement, with the UK more successful at 34%.

This could be caused by project overruns being more common at North American banks: only 12% said their projects had run to time and budget, compared to 40% of UK respondents. It could also reflect though the limitations placed on big data ambitions by the challenges surrounding the use of sensitive data.

Charts 8-13 offer further insights into what firms believe could be possible were they able to further exploit technology to protect data privacy. Asked to identify the big data models of most interest to them for the future, respondents registered the most excitement at the prospect of collaborating over data sets internally: 94% of respondents said this was an opportunity, with 84% confirming they could do this in future if data privacy were protected by technology.

In second place came "improving business insights", pegged as an opportunity by 84% and by 81% as something that could be done in future were the technology in place to protect data privacy.

"Firms have a strong understanding of what a data breach would cost them: 52% agreed or strongly agreed with this statement, rising to 73% for UK respondents."

CHART 8

SHARING AND COLLABORATING OVER DATA SETS INSIDE THE BUSINESS TO DRIVE INNOVATION

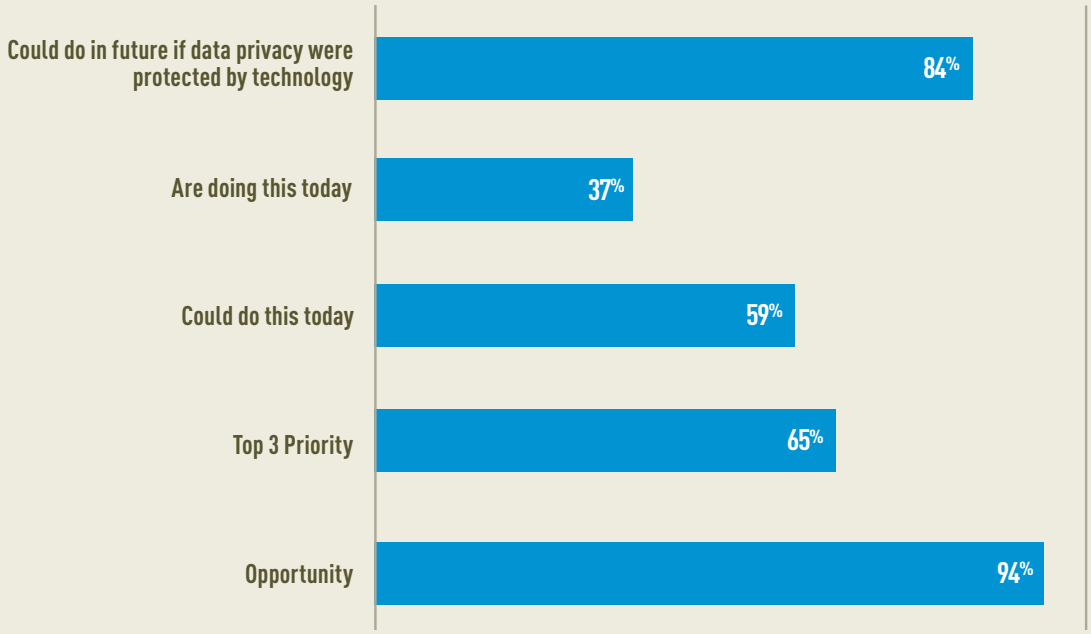


CHART 9

SHARING AND COLLABORATING OVER DATA SETS WITH THIRD PARTIES TO DRIVE INNOVATION

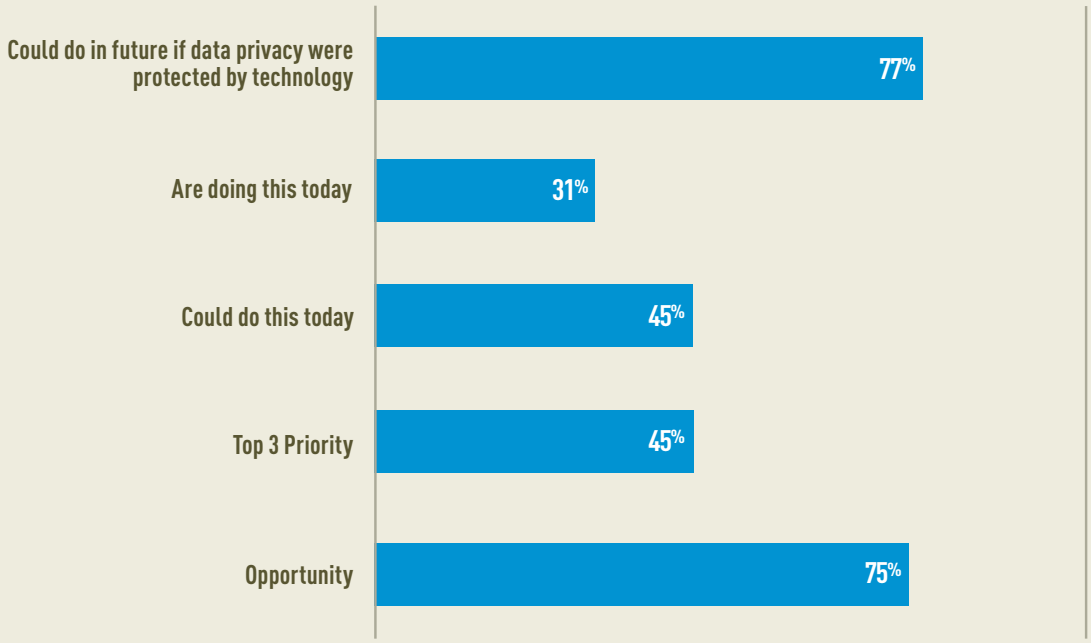


CHART 10

MOVING FROM A 'ONE SIZE FITS ALL' APPROACH TO TRUST WITH SENSITIVE DATA, TO A MODEL BASED ON LEVELS OF TRUST UNDERPINNED BY TECHNOLOGY CONTROLS

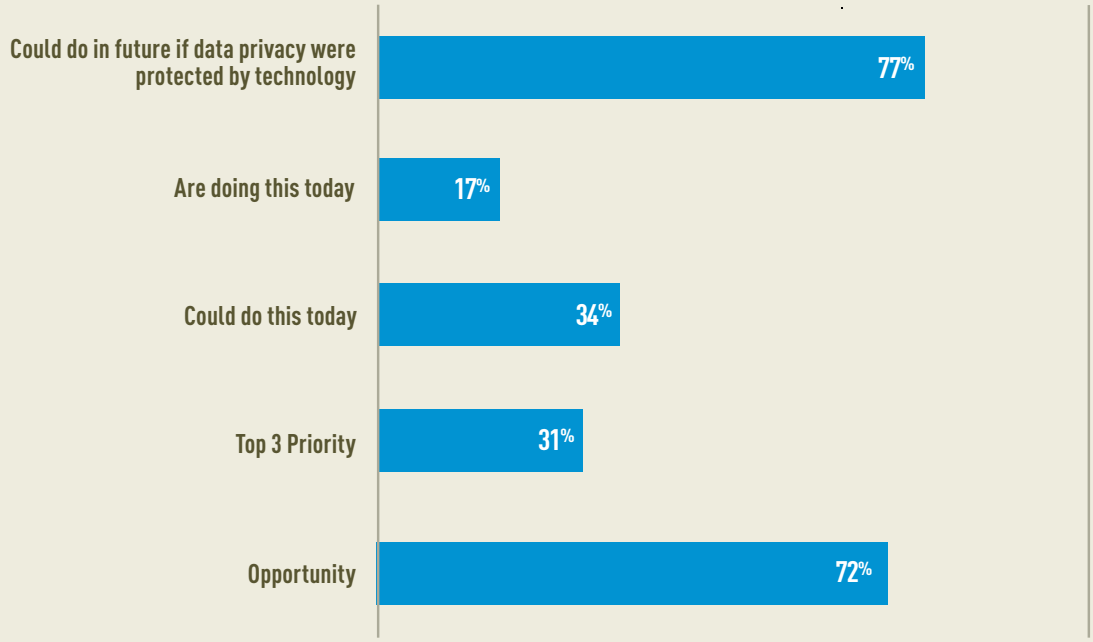


CHART 11

EXPLOITING CLOUD BY PUSHING WORKLOADS INTO THE PUBLIC CLOUD

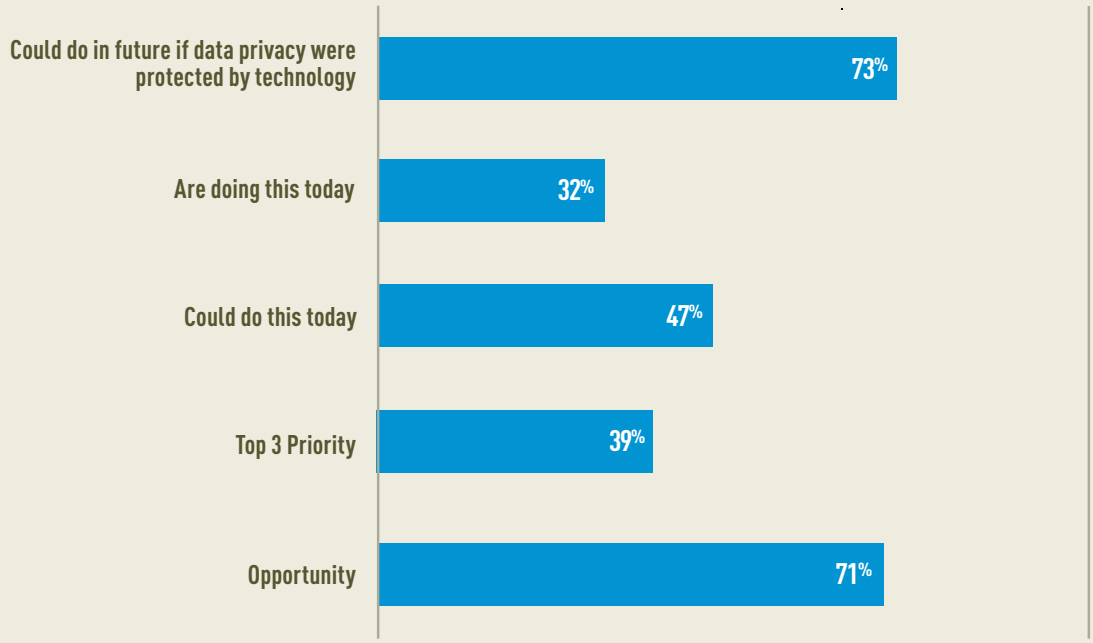


CHART 12

MAKING USERS RESPONSIBLE FOR THE ETHICAL USE OF DATA

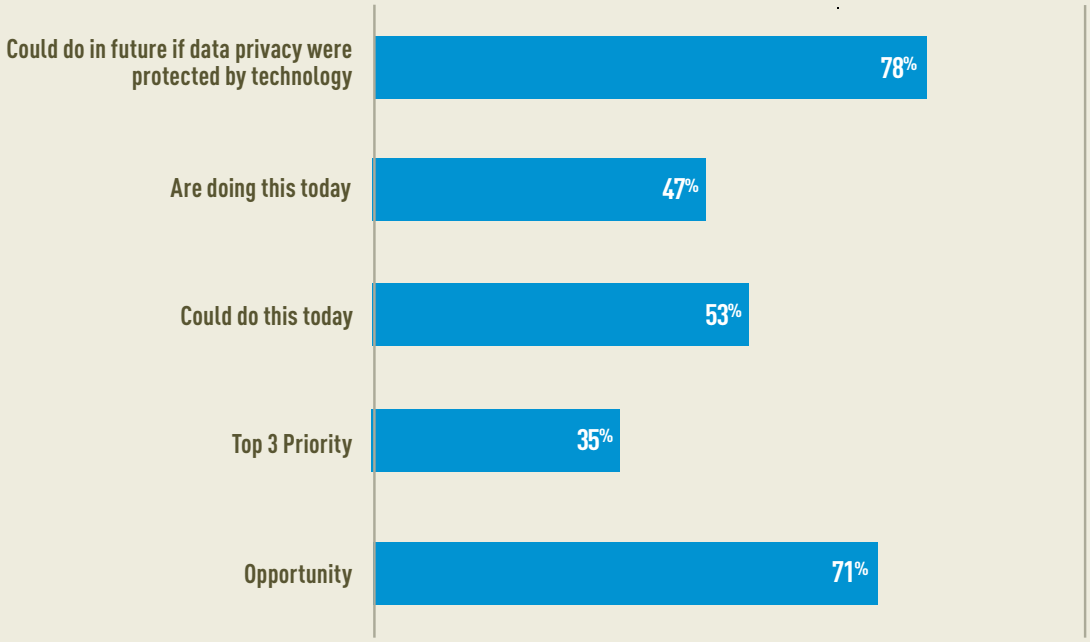
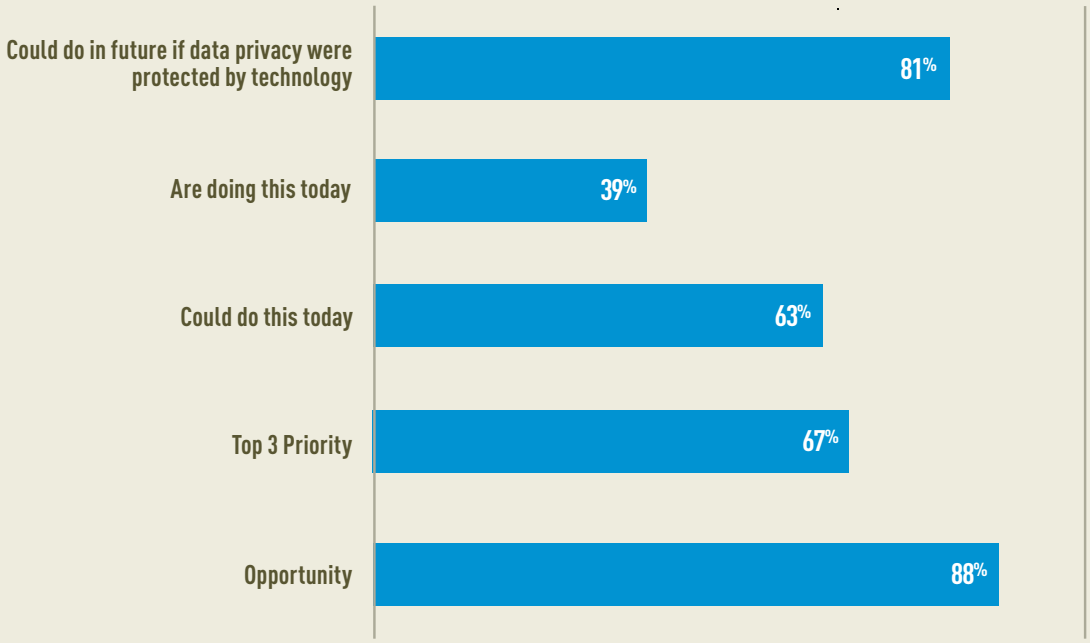


CHART 13

IMPROVE BUSINESS INSIGHTS (FOR EXAMPLE, RESEARCHING PRODUCT POPULARITY AMONG CUSTOMER SEGMENTS)



07
CONCLUSION

Finextra's five key takeaways from this research into the impact of data privacy challenges on the big data landscape for financial services firms are:

1) There is an **inherent tension** between firms' big data analytics ambitions and their data privacy obligations. Data privacy isn't a showstopper for big data, but it has to be considered in every big data project.

2) The data protection landscape is getting **more complex and challenging** for financial services firms to navigate. GDPR may be intended to bring harmony in Europe but it also adds new obligations and increased sanctions – and for global firms, the variations in data privacy laws country to country create another layer of complexity. Conflicting compliance requirements for GDPR and other regulations around KYC for example are also a problem.

   An extra dimension to this challenge comes from customers themselves, with their increased awareness of the value of their data and heightened desire for it to be protected.

3) The barriers – internal and external – that data protection and privacy put in the way of firms' realizing their big data analytics ambitions notwithstanding, the smart money is on viewing all these challenges as **opportunities**: opportunities to increase efficiency and to do more with data on the back of implementing privacy by design.

4) **Privacy by design** means thinking at the outset of every data-based project what the impact on personal, sensitive data might be, planning to mitigate that impact, and only ever gathering, storing or processing data that is actually needed.

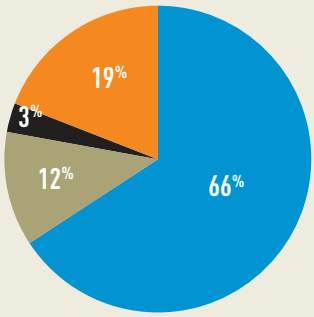
5) Careful, thought-through application of various technologies to enable **pseudonymization and/or anonymization** is a critical component of implementing privacy by design. There is no silver bullet and the risks – including that of re-identification – must always be carefully assessed, but used properly, privacy engineering technologies have a strong role to play in helping financial services firms to do more with their data.

   These techniques can empower financial institutions to strike the right balance between creating exciting new propositions to generate revenues and deliver customer benefits on the one hand, and protecting sensitive data to comply with regulation and retain customer trust on the other.

> "GDPR may be intended to bring harmony in Europe but it also adds new obligations and increased sanctions – and for global firms, the variations in data privacy laws country to country create another layer of complexity. Conflicting compliance requirements for GDPR and other regulations around KYC for example are also a problem."

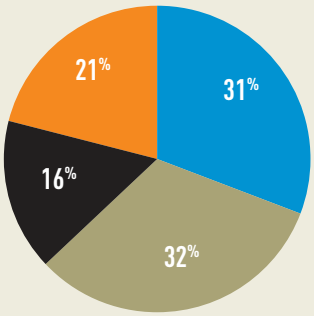
08
APPENDIX A – ABOUT THE SURVEY RESPONDENTS

TEAM SIZE

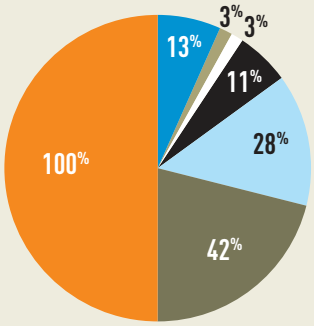


DEPARTMENT ANNUAL BUDGET SIZE

(51% of those surveyed responded)



RESPONSES BY REGION



09 ABOUT

Finextra

This report is published by Finextra Research.

Finextra Research is the world's leading specialist financial technology (fintech) news and information source. Finextra offers over 100,000 fintech news, features and TV content items to visitors to www.finextra.com.

Founded in 1999, Finextra Research covers all aspects of financial technology innovation and operation involving banks, institutions and vendor organisations within the wholesale and retail banking, payments and cards sectors worldwide.

Finextra's unique global community consists of over 30,000 fintech professionals working inside banks and financial institutions, specialist fintech application and service providers, consulting organizations and mainstream technology providers. The Finextra community actively participate in posting their opinions and comments on the evolution of fintech. In addition, they contribute information and data to Finextra surveys and reports.

For more information:
Visit www.finextra.com, follow @finextra,
contact contact@finextra.com
or call +44 (0)20 3100 3670

Privitar

Privitar is a London based software company that enables organizations to use, share and derive insights from data safely. Using patented privacy engineering technologies, Privitar products create opportunities by allowing broader use of valuable information assets for collaboration and sharing, while reducing the risk associated with storing, processing and using sensitive data.

Founded in 2014, Privitar is working with leading European and global financial institutions to help navigate privacy risks and the regulatory environment, while balancing the need to extract maximum value from sensitive data.

For more information:
Visit www.privitar.com or follow @privitarglobal
Contact info@privitar.com
Or call +44 203 282 7136

Finextra

Finextra Research Ltd
1 Gresham Street
London
EC2V 7BX
United Kingdom

Telephone
+44 (0)20 3100 3670

Email
contact@finextra.com

Web
www.finextra.com