Storage Anywhere and Everywhere

Dealing with the challenges of data fragmentation

Freeform Dynamics Ltd, June 2013

Gone are the days when one or two central systems could meet all of the organization's IT needs. We now have more departmental applications, collaboration and sharing environments, and even cloud-based services in the mix, with access via desktops, laptops, tablets and smartphones that can each store data in their own right. And to complicate things further, consumerization is putting more power into the hands of users to choose the equipment and facilities they use to get their job done. Against this background, are we in danger of losing control of our business information?

Key Points

Information-related requirements and expectations continue to rise
A recent study of 300 mid-sized organizations across the USA and UK highlighted how much access to good information matters to the modern business, and how critical it is to protect the data that's important to us. In specific terms, the majority report an increasing emphasis on better exploiting information for decision-making purposes, with a corresponding focus on collaboration facilities, anytime, anywhere access, and measures to secure and manage information more effectively.

But data fragmentation makes it harder to exploit information, as well as creating additional risk
In addition to IT-managed systems, so called 'shadow IT' means a significant amount of data is now stored under the control of departments and workgroups. Individual users then hold data locally on PCs and mobile devices (including personal equipment), and in consumer cloud services. Difficulties finding what's needed, conflicting versions of the truth, visibility and compliance issues, and the anguish caused by loss or leakage of unprotected data all stem from the way in which storage has become fragmented.

The most prominent consequence of data fragmentation, however, is elevated cost and overhead
Most study participants say fragmentation translates directly to wasted storage capacity due to repeated over-provisioning, multiple copies of data driving up storage volumes and costs, and a significant amount of additional overhead to manage the complexity. With this in mind, the overwhelming majority (4 out of 5 respondents) recognize the problem or danger of storage costs running out of control.

Cloud storage is part of both the problem and the solution
The risk of uncontrolled use of consumer cloud services aggravating data fragmentation challenges and risks is acknowledged by most, with concern about this being as high as BYOD. However, many also believe that more structured adoption of cloud storage is a good way of achieving consolidation without sacrificing the pervasive access and convenience craved by users. The principle of cloud-based archives is also broadly appreciated as a way of managing historical data more efficiently and effectively.

Elite performers light the way to effective, safe and efficient use of business information
Within our study sample, 39 Elite performers (a small but significant group) provide valuable insights into how best to deal with the challenges. They highlight the importance of knowing your data and defining clear policies around the storage, retention and protection by information type. This in turn allows investment in technology, services and processes to be prioritized objectively, and the Elite place a big emphasis on exploiting cloud options here. However, the need to educate business stakeholders on the value and risk aspects of better data management, as well as cost, also comes through very strongly.

The study upon which this report is based was designed, executed and interpreted independently by Freeform Dynamics Ltd. Data was gathered via in-depth telephone interviews with 300 respondents from SMBs or Small Enterprises across the USA and UK. All participants were either responsible for, or intimately involved in, information-related policy, planning and decision-making.

Study sponsored by

mimecast

CONTENTS

Introduction

Where is your organization's business data held? This sounds like a straightforward question, and locations such as your accounting system, customer database, or collaboration environment might immediately come to mind as you think about answering it. But what about your email server, the hard drives of PCs and laptops, or the internal memory of mobile devices, not to mention USB sticks and cloud services such as Dropbox or similar?

The reality is that a plethora of storage options are available today to your employees, which creates the potential for business information to end up spread across a diverse range of systems, devices and services, which may or may not be centrally managed.

In this report, we look at the degree to which business information has become fragmented in many organizations and some of the issues arising from this. We also explore the benefits and practicalities of taking a more proactive and structured approach to dealing with distributed data.

In order to help with our discussion, we'll be drawing on the results of a recent research study based on feedback gathered during in-depth telephone interviews with representatives of 300 mid-sized organizations across the USA and UK (see Appendix A for more details). Along the way, we will touch on areas such as data protection, archiving and the role of cloud, with the ultimate aim of identifying the key high-level ingredients of effective, safe and efficient information storage and management.

A good place to start, however, is with a quick review of the environment in which many organizations operate, and some of the important dynamics that are changing protection requirements and end users' expectations. You may recognize some of these.

Powerful and relentless forces at work

It is often said that an organization's greatest asset is its people. In the modern workforce, this increasingly includes knowledge workers who consume, create, manipulate, and share information as an integral part of their job. We then have mobile workers with a need to access systems and data while travelling or operating in the field. Activity here has been encouraged by the advances we have seen in telecommunications and mobile technology. Going hand in hand with all this is the move within many organizations to empower users so they can operate in a self-sufficient manner when assessing situations, making decisions and taking appropriate action (Figure 1).

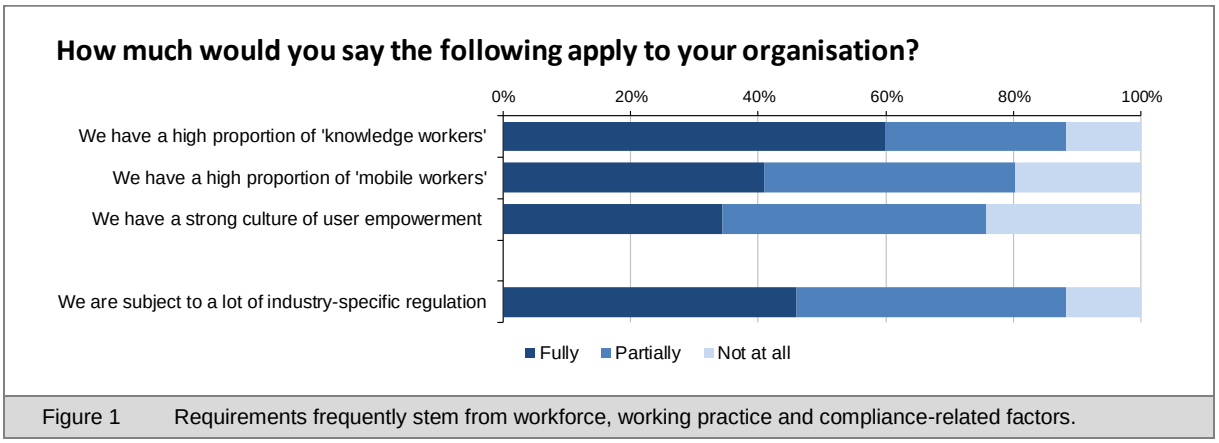


Figure 1 Requirements frequently stem from workforce, working practice and compliance-related factors.

From the chart presented here, we can also see that most organizations need to deal with at least some level of industry-specific regulation. This may be to do with health and safety, quality assurance, financial management, data privacy, or any other areas of your business subject to legislation.

These workforce, working practice and compliance-related factors all translate to a need to capture, store, manage and access information effectively, safely and efficiently. Demanding enough if you take a point-in-time snapshot of what's required in any given area, but when you overlay on this the relentless growth we see across the board in the diversity and volume of data building up within the business, it's not surprising that most report a range of escalating requirements and expectations (Figure 2).

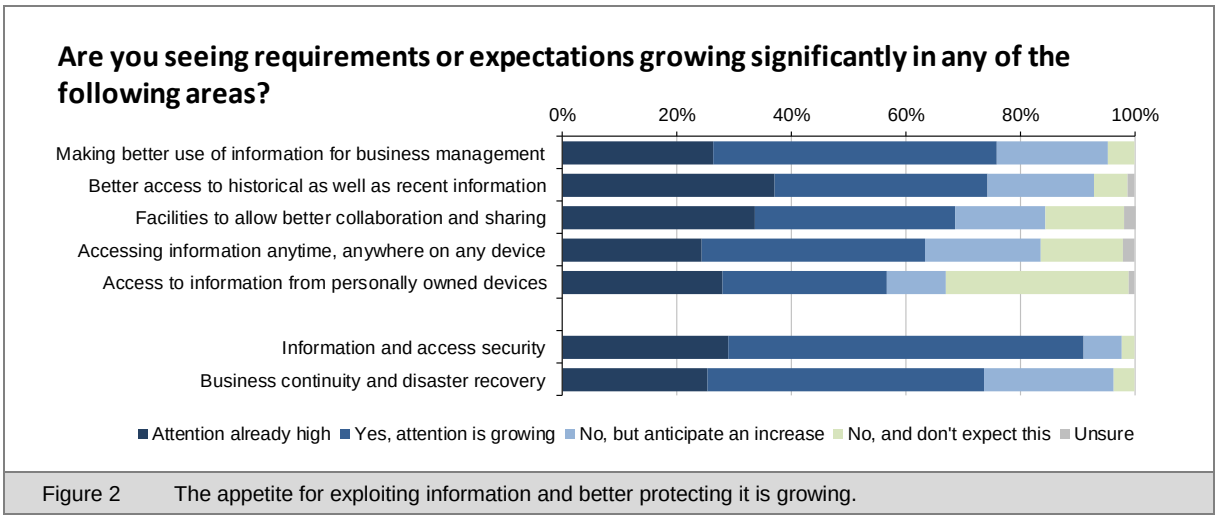


Figure 2 The appetite for exploiting information and better protecting it is growing.

Looking at this list, it is clear that the appetite for exploiting information for management and decision-making is growing (if attention is not already high), and the same is true when it comes to security and data protection. However, while requirements are escalating, so too are the challenges that stem from the way in which data is becoming increasingly more fragmented.

The inevitable fragmentation of storage

Respondents in our study unsurprisingly indicate that a great deal of critical and/or sensitive information is held in core application databases, along with file shares and collaboration environments deployed and supported by the IT team. Perhaps more interesting is the degree to which the company email system has become a key information repository, and the level to which IT departments are rolling out access to storage in the cloud (Figure 3).

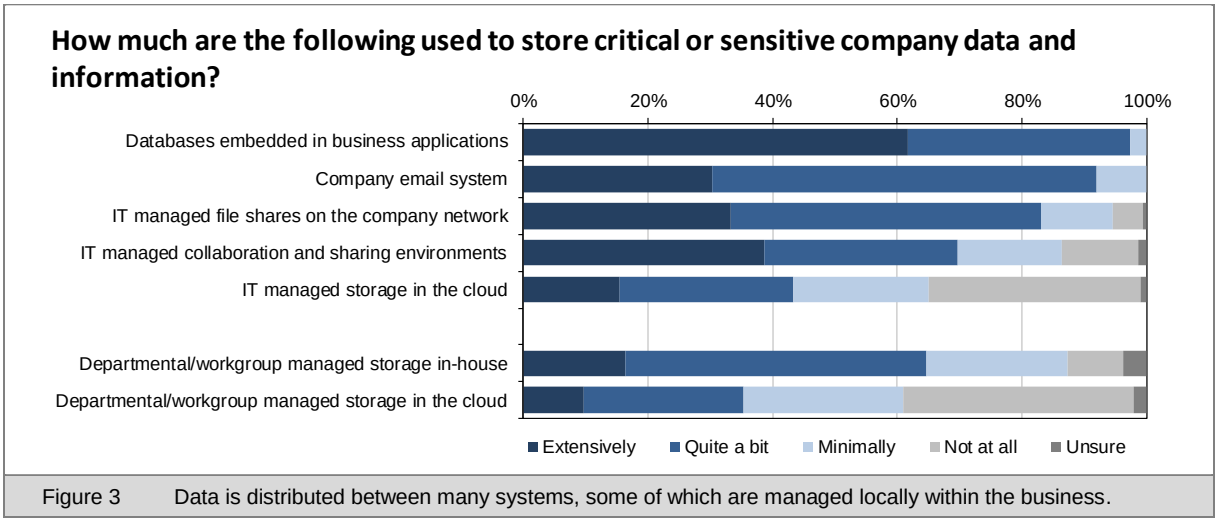


Figure 3 Data is distributed between many systems, some of which are managed locally within the business.

The other major observation from this chart is the incidence of what some are referring to as 'shadow IT', in which elements of information technology, in this case storage, are managed locally within a business department or workgroup. As part of this phenomenon, it is also notable that about a third report a significant amount of unilateral adoption of cloud services, in which critical or sensitive data is being held.

Already from this picture, the degree to which business information is typically distributed across multiple systems and services is apparent, but this is only the tip of the fragmentation iceberg. When we turn our attention to the activities of individual users, it's clear that business information is often spread even more widely (Figure 4).

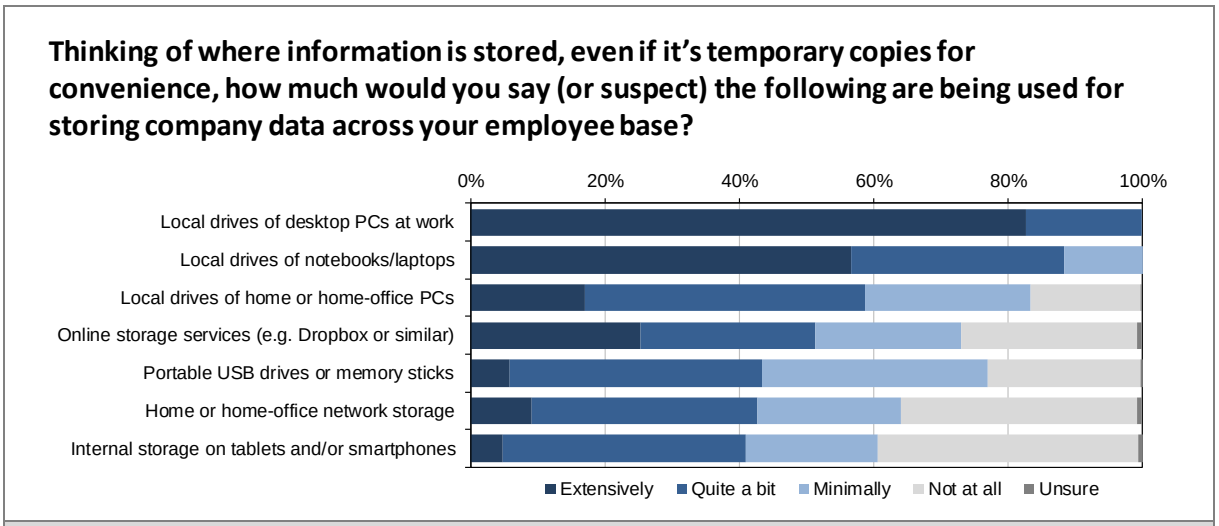


Figure 4 End user activity gives rise to yet more fragmentation, including across personal devices and stores.

In line with a lot of what we read in the press, we see many respondents reporting quite a bit of business data being held on smartphones, tablets and removable storage devices, or in cloud services signed up to by individual employees. We can expect such activity to continue to grow over time given the accessibility and convenience of both sophisticated mobile equipment and online storage offerings.

But we should not overlook the contents of company desktop and laptop PCs, which still account for the bulk of the personal and local storage we see. Given that pretty much any machine less than three or four years old will probably have 100Gb or more of internal disk space, which enables fast and immediate access to data, it is possible for users to accumulate a great deal of locally held information, whether endorsed by the organization or not.

As part of this, there is again significant evidence of email systems being used as repositories for business information in general, this time in a personal storage context in the form of off-line stores such as Microsoft Outlook *.PST files (Figure 5).

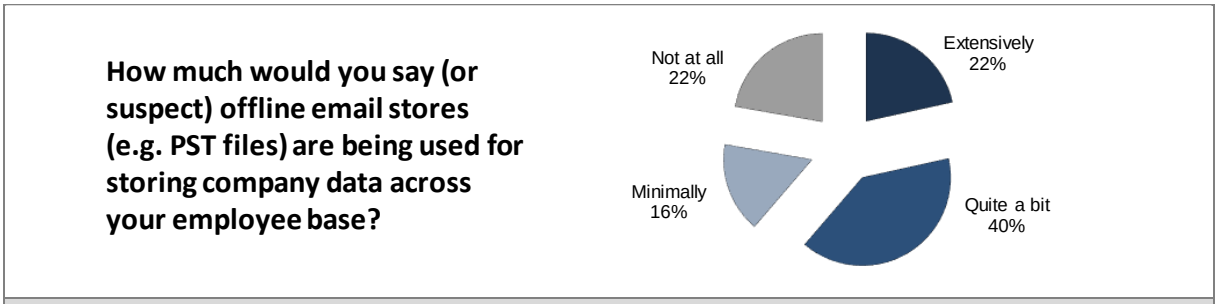


Figure 5 Offline email stores are often used as general repositories for business information.

This kind of behavior is understandable. With some of the core systems fragmentation we have seen, it's often easier for a user to search their local email 'archive' for a document they know was attached to a message, than to figure out which network share or intranet team site it is held on. This brings us conveniently to the question of how the activity we have been discussing impacts the business.

The triple-hit impact of data fragmentation

When we think of the way in which information is proliferating across many different types of storage, including mobile devices and cloud services, the first thing that springs to many people's minds is security, and this can indeed be a problem. However, fragmentation of business data gives rise to a whole range of other issues, and when we look at these, challenges are evident across three important business categories – the undermining of business value, the creation of unnecessary cost, and the escalation of risk (Figure 6).

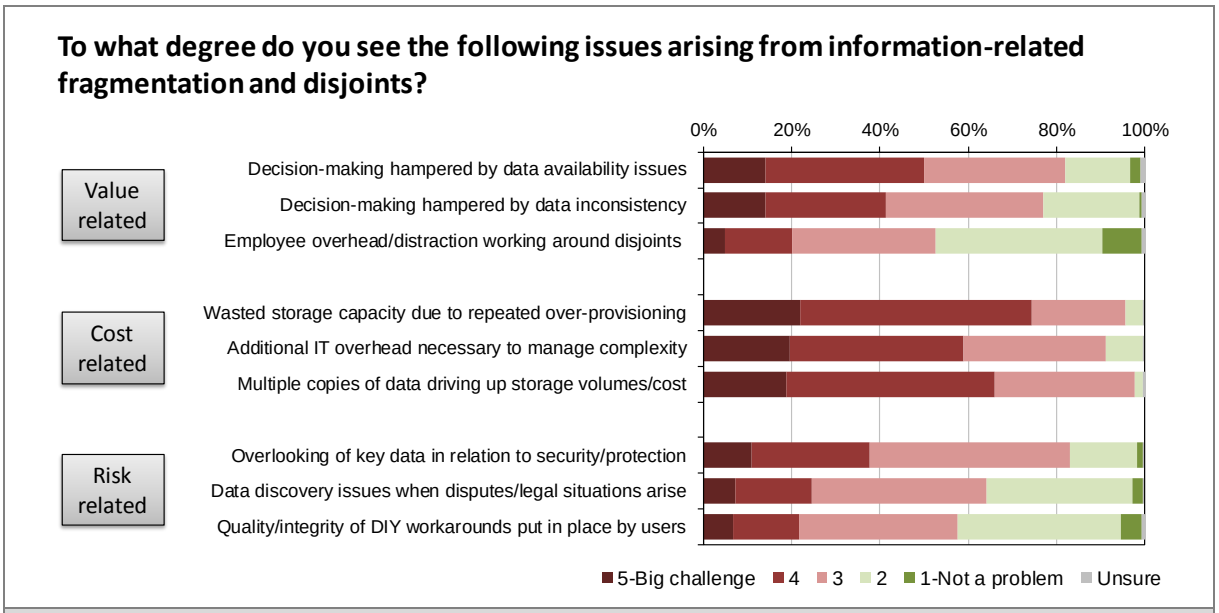


Figure 6 Cost-related pain of fragmentation is acutely felt, though risk and value related issues are evident.

What's interesting about this chart is that the pain of fragmentation seems to be most intensely felt in relation to cost and overhead, with significant, but noticeably less emphasis on the risk-related issues that often gain most attention in the press. This reflects the fact that IT professionals take risk very seriously (despite the fact that users often don't) and do what they can to mitigate it, even if it means putting a lot more effort into dealing with the complexity and working around the constraints that exist.

Having said this, there is a sense that keeping up with some of the dynamics we mentioned at the outset is becoming harder, and the pressure on IT departments with limited resources and budgets is clearly beginning to show (Figure 7).

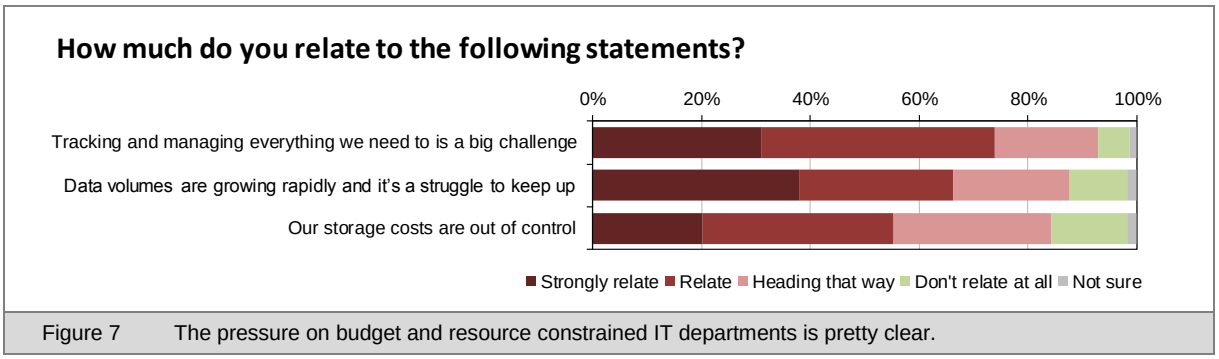


Figure 7 The pressure on budget and resource constrained IT departments is pretty clear.

Unfortunately, the way in which user behavior is evolving often adds to this pressure. At the time of writing for example, the practice of BYOD ('Bring Your Own Device'), in which personally owned equipment is used for work purposes, is stimulating a lot of debate. The cost and risk implications of granting users more freedom in relation to the technology they use are particular concerns.

Added to this, users are increasingly setting up personal cloud service accounts and using them to create, store and share business data. Indeed, it is not uncommon for even a single user to store data in multiple public clouds. They may use something like Dropbox to coordinate document access across various devices, a service like Evernote to capture discussions during business meetings, and a cloud-based office service for convenient document editing in a Web browser.

Then, just for good measure, they back-up or sync their iPhone, Android device or Windows phone with the relevant cloud service from Apple, Google or Microsoft, including business contacts, diary entries, notes and tasks, all of which may contain confidential information.

Multiply such behavior across tens or hundreds of employees, and it's no wonder many respondents in our study are as concerned about the use of personal cloud services as they are about storing data on personal devices (Figure 8).

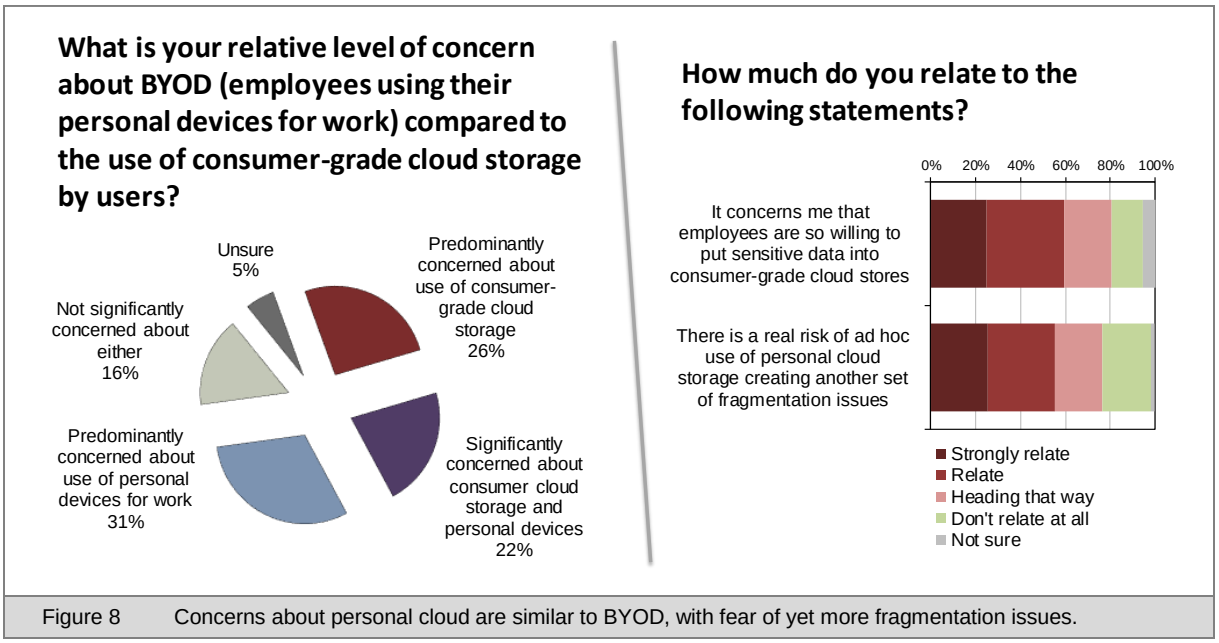


Figure 8 Concerns about personal cloud are similar to BYOD, with fear of yet more fragmentation issues.

Beyond the concerns about security, many also recognize the danger of personal cloud storage creating a whole new set of fragmentation issues, aggravating the problems previously identified.

Learning from those who are getting results

When looking across the results of the study, it is evident that some organizations do much better than others when it comes to exploiting business information and managing the costs and risks associated with this. But what is it that those achieving the best results are doing that others are not?

We can get a feel for this by grouping respondents based on their aggregate performance across the areas of effectiveness, risk management and efficiency (Figure 9).

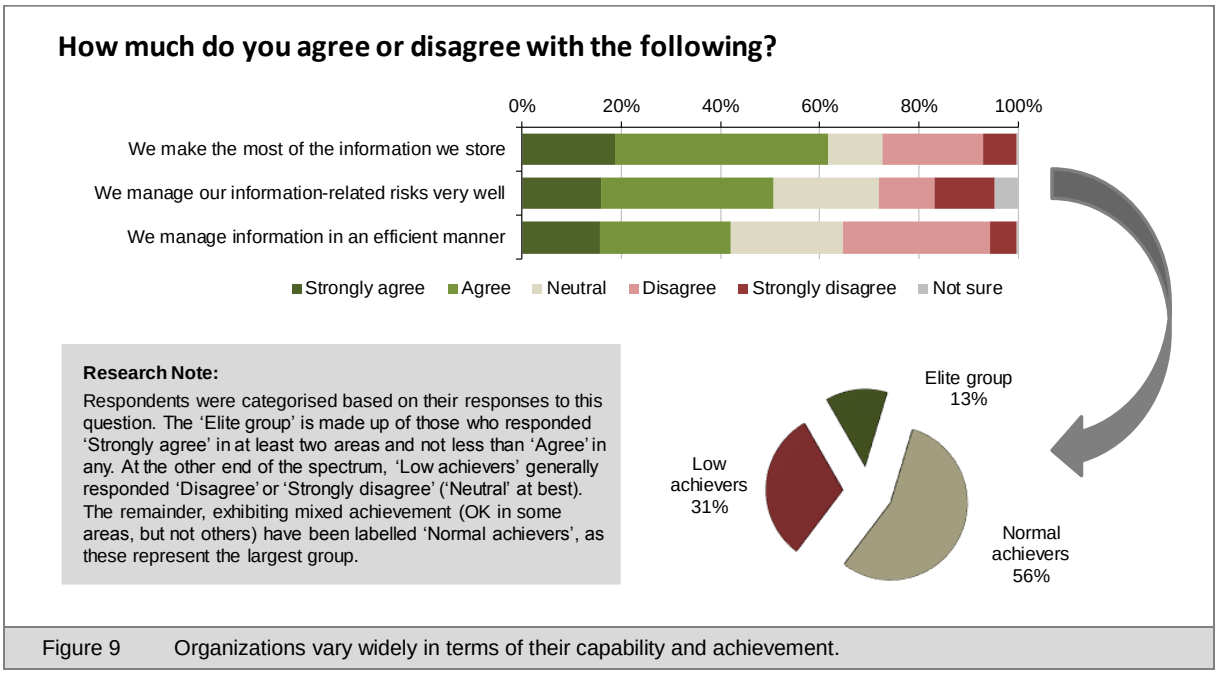


Figure 9 Organizations vary widely in terms of their capability and achievement.

If you are interested in comparing how well information is managed overall by organization size, industry and geography, we have provided some data and additional notes in Appendix B.

In the meantime, let's look at what we can learn from those achieving the best results, particularly the 'Elite' group of respondents (numbering 39 out of the total sample of 300) who get the most from their data, manage risks effectively, and control costs well.

Knowing your data and defining the right policies

At the most basic level, one of the things that sets those in the Elite group apart from others is the degree to which they understand their data. As an example, without exception, they each have comprehensive policies and procedures in place to classify data so they can make objective decisions on what needs to be kept and for how long (Figure 10).

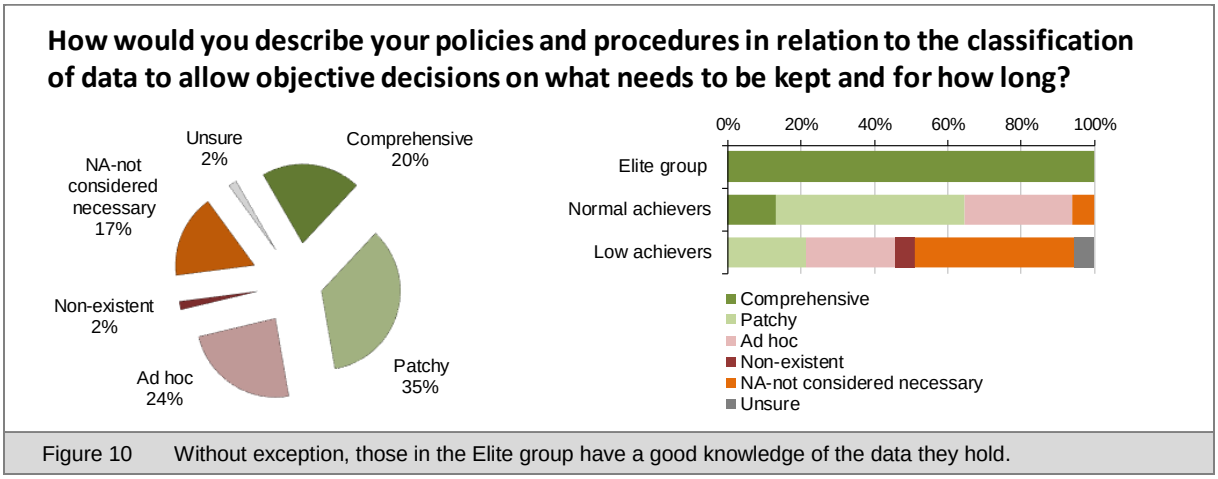


Figure 10 Without exception, those in the Elite group have a good knowledge of the data they hold.

Apart from being able to reduce the amount of information that needs protecting and managing (because you can dispose of unnecessary data with confidence), if you have a good feel for the different types of information you hold, you can more easily judge what should and shouldn't be held where. In line with this, those in the Elite group are much more likely to have comprehensive policies in place with regard to approved storage mechanisms and how they are utilized (Figure 11).

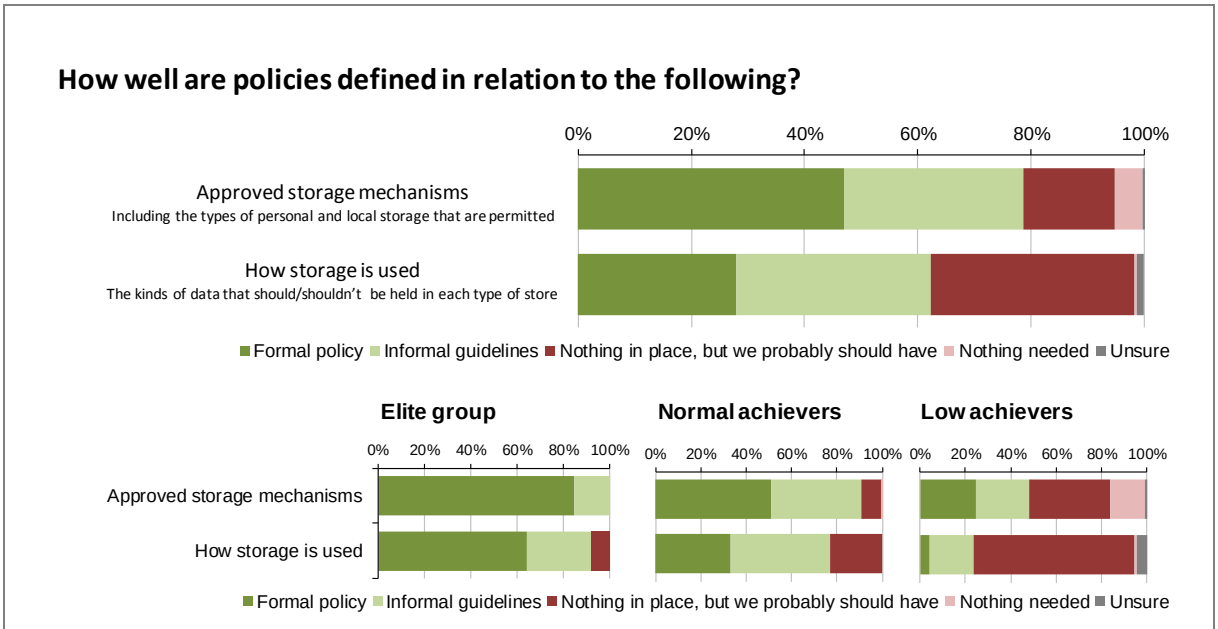


Figure 11 Elite performers are clear on which storage mechanisms are approved and how to use them.

At the other end of the spectrum, those in the Low Achiever category typically provide limited guidance to users on how to store and share data safely and efficiently, even though they know they should be doing more. This highlights a significant gap given that Low Achievers make up nearly a third of our sample.

When it comes to more specific areas of risk management, we again see the Elite being more likely to have appropriate policies or guidelines in place, while many others have little or nothing (Figure 12).

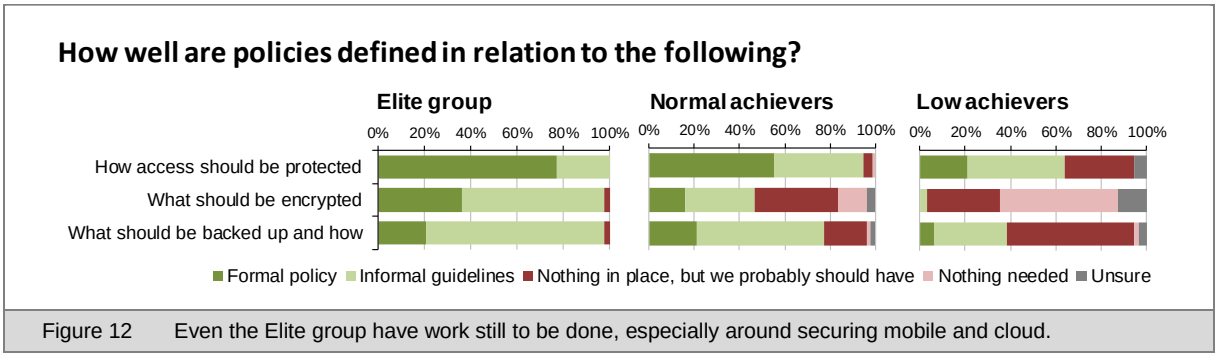


Figure 12 Even the Elite group have work still to be done, especially around securing mobile and cloud.

Looking more closely at this chart, it is telling that even those in the Elite group are relying more on informal guidelines at the moment when it comes to encryption and backup. Contributing to this is undoubtedly the fact that figuring out how to handle the rapidly evolving mobile device and cloud aspects of information storage is still a work in progress, even for those who are currently ahead of the game.

Putting the right systems and processes in place

Of course, defining what you should be doing is one thing, but enabling effective, safe, and efficient use of information typically also relies on putting the right kind of technology and processes in place. It will therefore come as no surprise that those in the Elite group are generally in much better shape when it comes to the implementation of specific measures to ensure adequate protection of data (Figure 13).

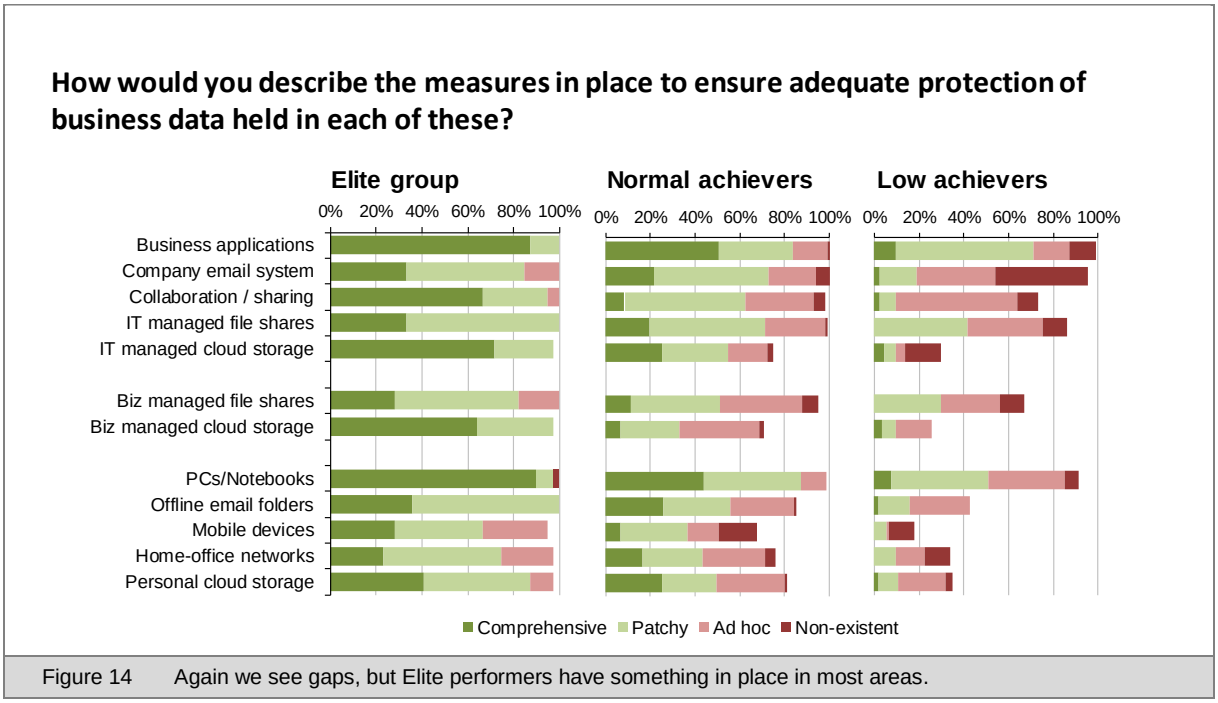


Figure 14 Again we see gaps, but Elite performers have something in place in most areas.

Elite performers indicate a degree of patchiness and work in progress here too, but again most notable are the capability gaps among Normal and Low achievers.

When looking at the above chart, it is also important to remember that the presence or absence of specific data protection measures does not just translate to a question of risk. Where solid systems and processes are in place, particularly when automation has been brought to bear on the problem, the cost and overhead of routine monitoring, management, troubleshooting and support can be lowered considerably. Furthermore, the substantial expense and distraction of dealing with crises and taking remedial action when things go horribly wrong can more easily be avoided.

If, like many of the participants in our study, you know that you have significant gaps in your data protection regime, it is worth exploring areas like data replication, snapshotting, image-based backup, even continuous data protection (CDP) and high availability (HA) architectures. If these terms are unfamiliar, you'll find some useful pointers in our report entitled "Enabling Rapid and Effective IT Recovery: Insights and tips for small and mid-sized businesses" which can be found on www.freeformdynamics.com.

We would also recommend exploring what's going on in the market in relation to end-point security and mobile device management (MDM), which are both hot areas where solutions are evolving rapidly. These can help with protection and management of data held on PCs, Macs, tablets and smartphones.

Meanwhile, in addition to comprehensive data protection, those in the Elite group are also much more likely to have archiving systems and procedures in place to move information to long-term, read-only storage when it is no longer active (Figure 15).

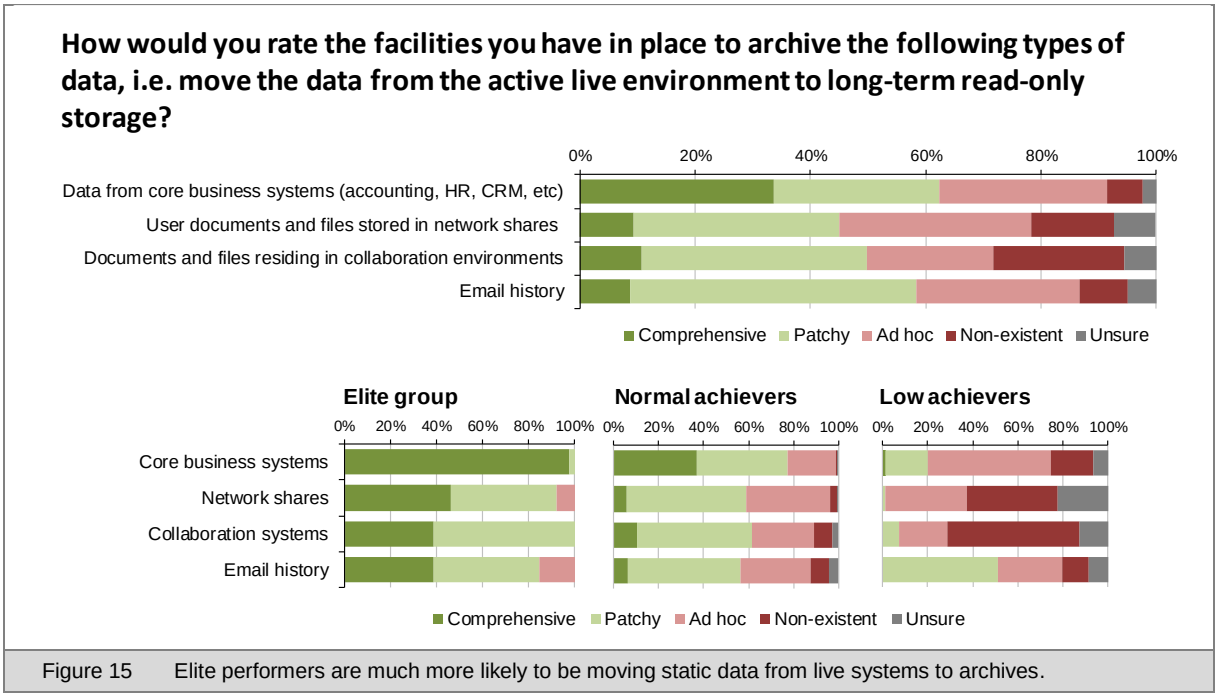


Figure 15 Elite performers are much more likely to be moving static data from live systems to archives.

The logic here is not difficult to work out. Data held in "live stores" that are continuously changing needs to be backed up, held on high performance media if speed of access is important, and replicated or mirrored if a degree of resilience is required. All of this consumes budget and resource, as well as impacting disaster recovery times should a system fail in a way that requires data to be restored (the more data you have in a system, the longer it will take to reload).

By contrast, archived data is by definition unchanging and accessed less frequently, so can live on lower performing and therefore lower cost storage. It also obviously requires a lot less administration from a systems perspective, and has an additional advantage from a compliance perspective because archived data is beyond the ability of users to tamper with or inadvertently delete.

So why, given these advantages, are most organizations not currently taking full advantage of data archiving, or not using it at all? Some respondents say it's because archiving is viewed as more trouble than it's worth, while others allude to difficulties getting business people to agree on what can and can't be archived. The most common blocker, however, is a perception that historical data needs to stay in the live environment to remain accessible, which is clearly not the case with modern archiving solutions.

Nevertheless, concerns about the cost, difficulty and access implications of archiving brings us onto the potential role of hosted services.

Exploiting the cloud and hosted services

As we saw earlier, ad hoc or uncontrolled use of cloud storage is part of the problem that many are trying to manage. Cloud services, however, can also be part of the solution.

Structured and properly supported adoption of cloud storage can be used, for example, as a way of consolidating data for easier protection and management. This is in principle no different to consolidating data using an in-house storage infrastructure or collaboration environment, but many who have gone down the in-house route have found that users still turn to personal cloud storage because it's more convenient. The advantage of consolidating storage in the cloud, particularly when facilities for synchronization with PCs, tablets and smartphones are included, is that it preserves the anytime, anywhere access benefits that users have come to value, of course with appropriate controls applied.

Picking up on our discussion of archiving, this represents another way of potentially exploiting cloud services. Hosted solutions can take much of the pain and cost out of implementation, and again a big benefit is that convenient access (subject to policy) is preserved. In fact, an 'always online' cloud service that users can get to easily via a browser or mobile app feels a lot less like an archive and a lot more like an interactive repository of historical business data – a subtle but very significant difference.

The value of exploiting cloud computing in the ways we have outlined is corroborated by the level of adoption and appreciation by those in the Elite performance group (Figure 16).

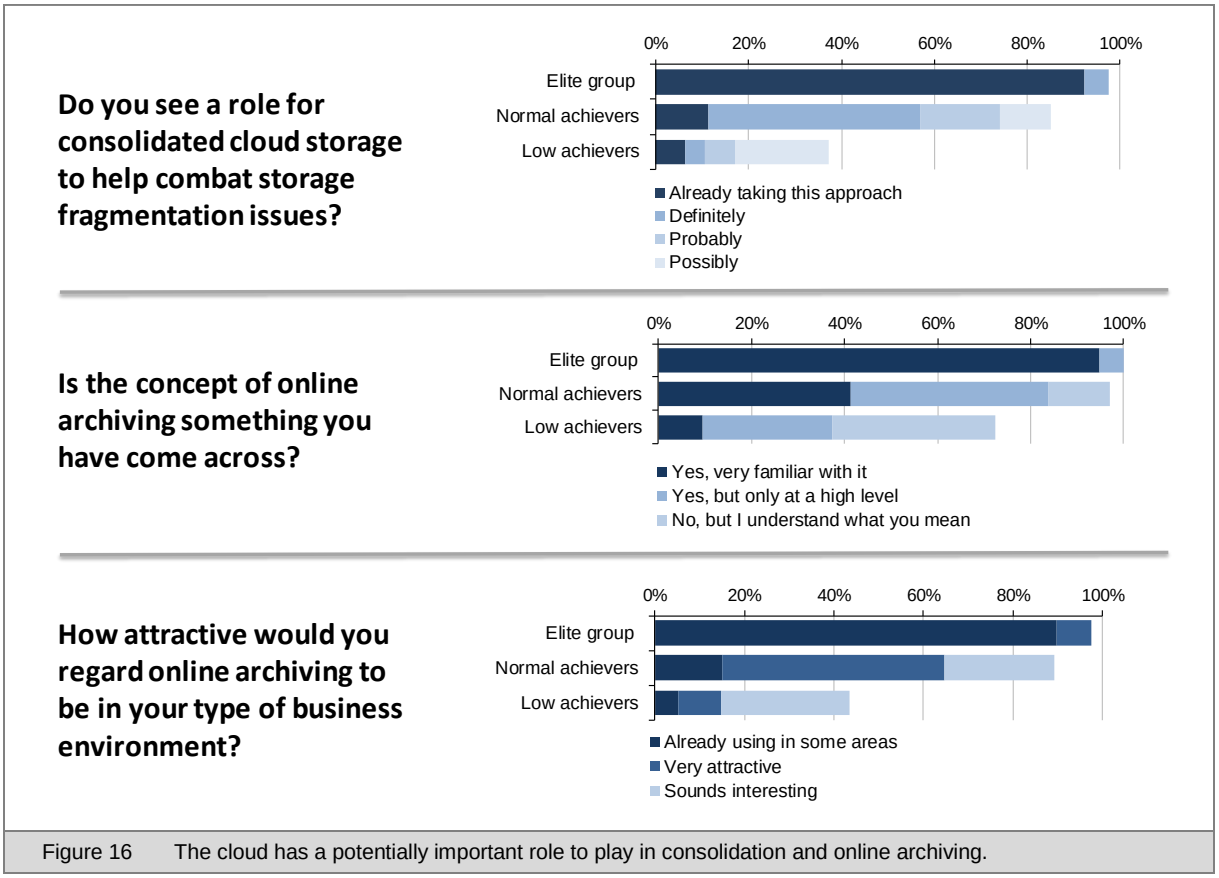


Figure 16 The cloud has a potentially important role to play in consolidation and online archiving.

So, we are building a good picture of the practices and solutions that can make a significant difference to how well the problems of data growth and fragmentation can be managed. But what are the practicalities of driving improvement?

Implementing change for the better

When scoping improvement activity, you need to avoid trying to boil the ocean, especially if your organization looks more like those in the Low Achiever group at the moment. We would therefore suggest beginning with a quick, open and honest review of current capability to identify areas of priority. For some there will be gaps in relation to core business systems that should probably be dealt with first, but for others it could be tackling fragmentation-related issues arising from end user activity.

At some point, you will need to perform at least a basic classification of the data you hold, if you haven't done so already. Without this, it's pretty much impossible to make decisions on the kind of retention, protection and storage policies to implement. You can then start looking at technology and services in the areas you have prioritized, including cloud options where appropriate.

Of course, all of this requires an investment in time, resources and money, which needs to be justified. As we have seen, many respondents in our study judge the cost impact of fragmentation to be significant, which means that in theory it should be possible to make the case based primarily on cost savings. The challenge here is collating the relevant data; you might instinctively know that investment in storage management and archiving solutions will save the organization money in the long run, but can you prove it? IT vendors and service providers can often help here with models and templates (so you at least know the kinds of costs you should be considering), but regardless of how well your ROI arithmetic comes out, indications are that broadening the business case beyond cost is useful.

If we look at what our Elite performers say motivates business executives, for example, we see value, user productivity and risk related drivers dominating the top half of the list, with money saving towards the bottom (Figure 17).

Figure 17 It helps if business stakeholders appreciate value and risk, as well as the cost perspective.

While those in the Low Achiever group may be there partly because of lack of attention at executive level, when it comes to investment in IT, particularly in relation to infrastructure and tools, the biggest impediment to securing management buy-in is usually simply a lack of understanding. A degree of education may therefore be required to help executives connect the dots between lack of business visibility, employee productivity issues, frequent data related crises or security breaches, and under-investment in information management.

Pulling it all together

We have covered a lot of ground in this report, and based on the results of our study have seen that growth in data volumes and diversity, coupled with IT-related trends towards departmental autonomy and end user empowerment, has created a data fragmentation problem. This in turn is preventing most organizations from making the most of their business data, as well as elevating both cost and risk.

By looking at the behavior and practices of an elite group of 39 respondents who are achieving particularly good results, however, we are able to identify the ingredients of success (Figure 18).

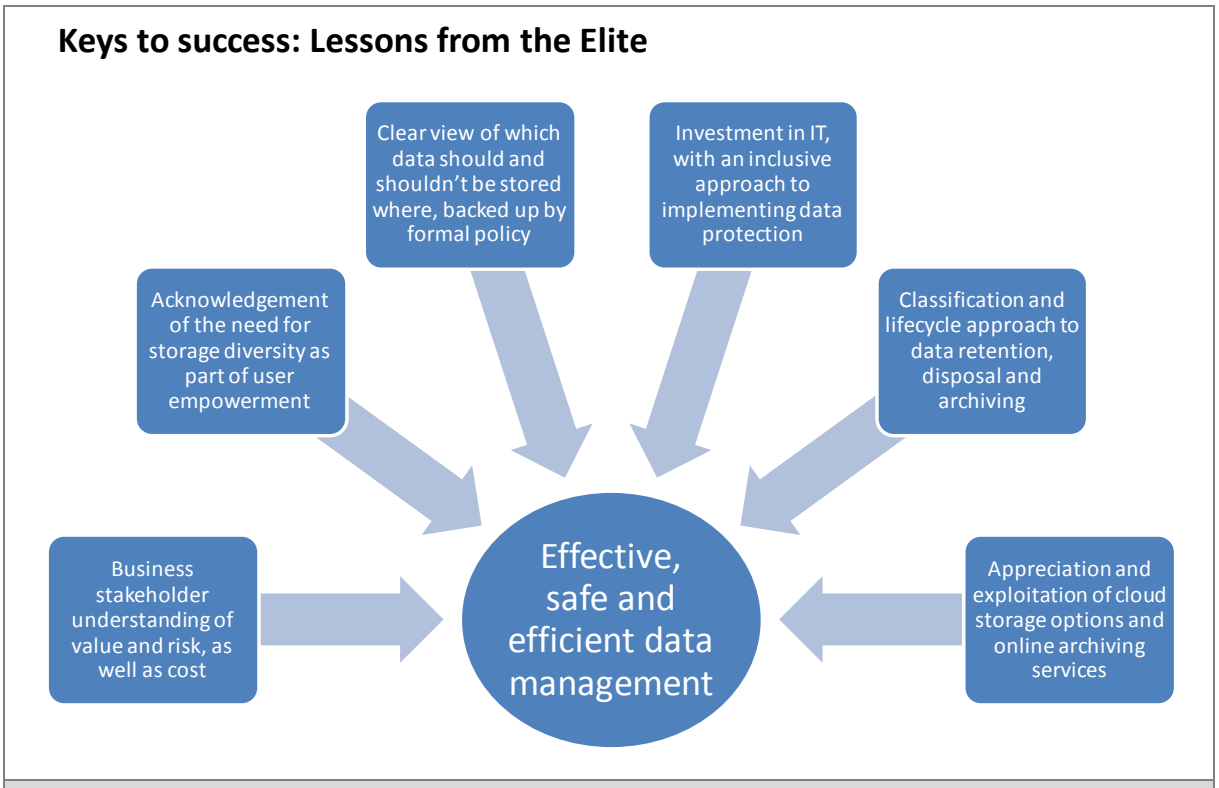


Figure 20 A range of behaviors, approaches and solutions correlate directly with better performance.

When looking to drive improvements, the journey starts with getting business stakeholders to recognize data management and protection as primarily a business issue rather than an IT concern, and to appreciate its importance from a value and risk, as well as a cost perspective.

Being realistic, while those in IT might fantasize about a world in which all data is centrally held and managed, for the vast majority of organizations this is neither practical nor desirable from a business perspective. Browser-based interfaces to core business applications and desktop virtualization techniques may help to centralize storage for some types of user, but we are a long way off having the pervasive, high-speed, low-latency connectivity that would avoid local storage across the board. In the meantime we need to accept storage diversity as a key requirement for user empowerment and productivity in most environments.

With this in mind, one of the main things that sets the Elite group apart from others is not whether data is spread across multiple locations, but how well it is managed. Inclusive backup and controlled synchronization, along with automated archiving into centralized repositories, are useful techniques here – it's the difference between distributed storage and storage fragmentation. Knowing your data is clearly a pre-requisite for taking such a managed approach, as is an inclusive set of measures that deals with all sensitive or critical data, and all of the devices and cloud services used to hold it.

Finally, Elite performers take a lifecycle approach to defining requirements, which allows them to reduce costs and overheads by either disposing of or archiving data when the time is right. And to overcome natural user desire to keep historical data available, most are turning to some form of online archiving.

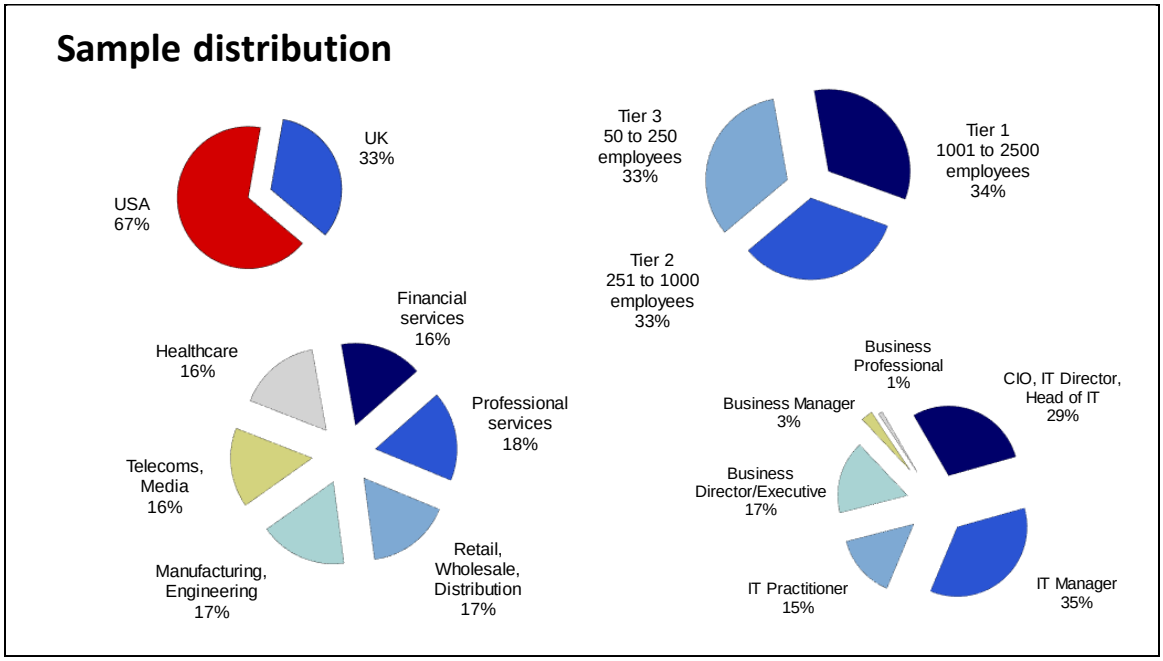
In order to help you make use of these insights from the best performers participating in our study, we have produced a more action-oriented 'Checklist for Success' which is presented in Appendix C.

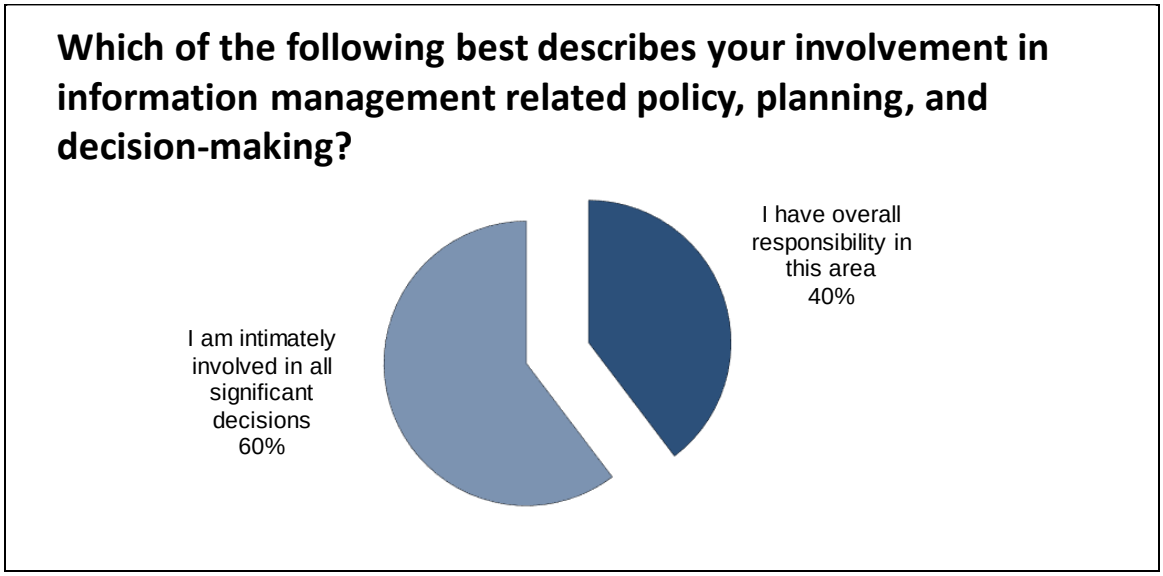
There's a lot more we could say on the topic of information management, but we hope our discussion here has at least helped to put the challenges of data fragmentation into perspective, and provide some ideas on how best to deal with them.

Appendix A: Study sample

The study upon which this report is based was designed, executed and interpreted on an independent and objective basis by Freeform Dynamics Ltd. Data was gathered via in-depth telephone interviews with 300 respondents from SMBs or Small Enterprises across the USA and UK. All of those participating were either responsible for, or intimately involved in, information related policy, planning and decision-making in their organization.

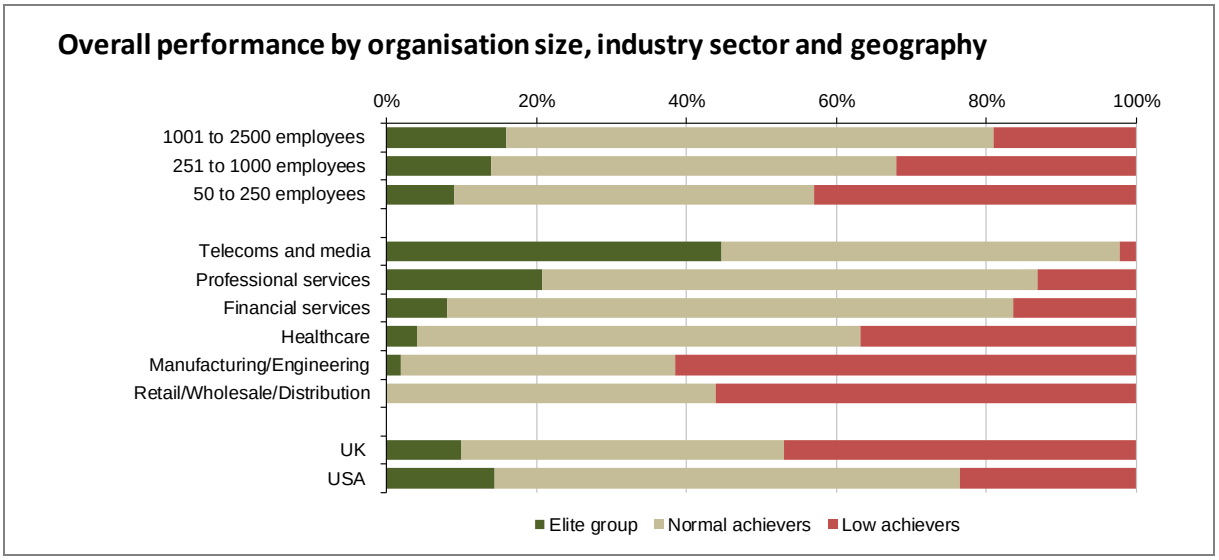
The sample distribution was as follows:





Appendix B: Analysis of overall performance by segment

Participants in the study were divided into capability/achievement-related groups based on a high-level self-assessment of how well they exploit, protect and manage their information. This allows us to look at differences in performance by organization size, industry and geography:



Larger organizations generally fare better than smaller ones, but the differences are not as great as we might expect; the luxury of more resources and more sophisticated IT is typically accompanied by increased scale and complexity of operations, so the two cancel each other out.

At the bottom of the chart, we see a difference between the UK and USA, the two geographies included in the study, but the reality is that it's probably not as great as the research results would suggest on face value. Experience tells us that Americans are more likely to emphasize the positives and downplay the negatives than the British when responding to questions, but it's impossible to adjust for this cultural factor in a study of this nature.

The biggest differences observed are based on industry sector, and these are not difficult to understand. Managing information and/or knowledge is a core business activity in industries such as telecoms and media, and professional services. The regulated nature of financial services and healthcare means information management must be at least 'adequate', while many industrial or goods-centric businesses can function without paying too much attention to anything other than the basics.

Perhaps the most significant observation from the above chart, however, is that regardless of how we cut the data, a spread of capability is evident in each sector. To put it another way, we see examples of good and poor achievement in every size, industry and country group. This is why we have based most of our discussion in this report on the attributes and behavior of the Elite versus other groups, as the insights arising are then relevant across the board.

Appendix C: Checklist for Success

Through analysis of the priorities and behavior of the Elite performance group, it is possible to distill out a list of action items for dealing with today's distributed information environment in an effective, efficient and safe manner:

1. **Measure ROI in business and IT terms** – the elites say their business executives take a balanced view, considering business visibility, enhanced decision-making and user productivity, as well as risk management.
2. **Understand your data** – establish comprehensive policies and procedures to inform decisions on what data is important and valuable, and what needs to be kept, where and for how long.
3. **Establish policies for how different storage mechanisms should be used** – define and communicate policy on what data should be held in enterprise, personal, network, cloud and local storage.
4. **Protect your data everywhere** – give clear guidance on data encryption and back up practices. Pay particular attention to mobile devices and cloud storage, especially the management of data on personal or 'non-authorized' devices and services.
5. **Protect all your data types** - implement specific measures to ensure adequate protection of all forms of data in your organization. From PCs, file shares and application data to IT managed or personal cloud storage and, crucially, your company email system.
6. **Streamline routine operations -** use systems, processes and automation to reduce the costs associated with monitoring, management, troubleshooting and support, and to avoid the expense and distraction of dealing with crises.
7. **Use cloud and hosted services** – consolidating storage in the cloud, particularly when synchronization with PCs, tablets and smartphones is included, preserves anytime, anywhere access benefits for users and you have control.
8. **Archive** – reduce storage costs by using archiving systems and procedures to move information to long-term, read-only storage when it is no longer active. Modern archiving means users can still get easy access to it when needed.
9. **Take archive to the cloud** – an 'always online' cloud service that gives users access to data via a browser or mobile app effectively turns an archive into a valuable interactive repository of historical business data.

It should be clear from this that there are no silver bullets in this space. Success is dependent on a proactive and organized approach based on the right blend of policy, process, technology and services.

About Freeform Dynamics

Freeform Dynamics is a research and analysis firm. We track and report on the business impact of developments in the IT and communications sectors.

As part of this, we use an innovative research methodology to gather feedback directly from those involved in IT strategy, planning, procurement and implementation. Our output is therefore grounded in real-world practicality for use by mainstream IT professionals.

For further information or to subscribe to the Freeform Dynamics free research service, please visit www.freeformdynamics.com or contact us via info@freeformdynamics.com.

About Mimecast

Mimecast (www.mimecast.com) delivers cloud-based email management for Microsoft Exchange, including archiving, continuity and security. By unifying disparate and fragmented email environments into one holistic solution that is always available from the cloud, Mimecast minimizes risk and reduces cost and complexity, while providing total end-to-end control of email. Founded in the United Kingdom in 2003, Mimecast serves more than 7,000 customers and 1.8 million users worldwide and has offices in Europe, North America, Africa, Asia Pacific and the Channel Islands.

For more information, please visit www.mimecast.com.

Terms of Use