computing
research

RESEARCH PAPER

GDPR is a state of mind, not just a technology solution

June 2017

Sponsored by ca technologies

CONTENTS

Executive summary

The European General Data Protection Regulation (GDPR) will be upon us on May 25th 2018, and contrary to enduring public opinion, Brexit won't make any difference. The GDPR comes in response to global shuffling of privacy laws to meet the growing demands of cloud, data security and other technological needs. The US Safe Harbor framework has been replaced with Privacy Shield, and on top of this is the e-Privacy Regulation, which takes specific interest in electronic communications, cookies for tracking user behaviour online, and other issues around personal data and consent.

While plenty of thought leadership and 'good sense' articles, and indeed white papers, already exist around the topic, the GDPR debate arguably lacks more focused discussions around techniques, tools and solutions that can be employed to integrate into an organisation's daily life.

To that end, *Computing* surveyed 115 UK business decision-makers in large enterprises of strictly 5000 or more employees – some with 15,000 or more – to find out their concerns about GDPR and how these concerns pertain to the everyday issues of storage, security and software development.

Introduction

GDPR is coming, and there is no black and white compliance solution. An organisation cannot simply deploy a single 'silver bullet' technology solution to manage it all, because preparation and governance involves the full gamut of people, processes, attitudes and behaviours.

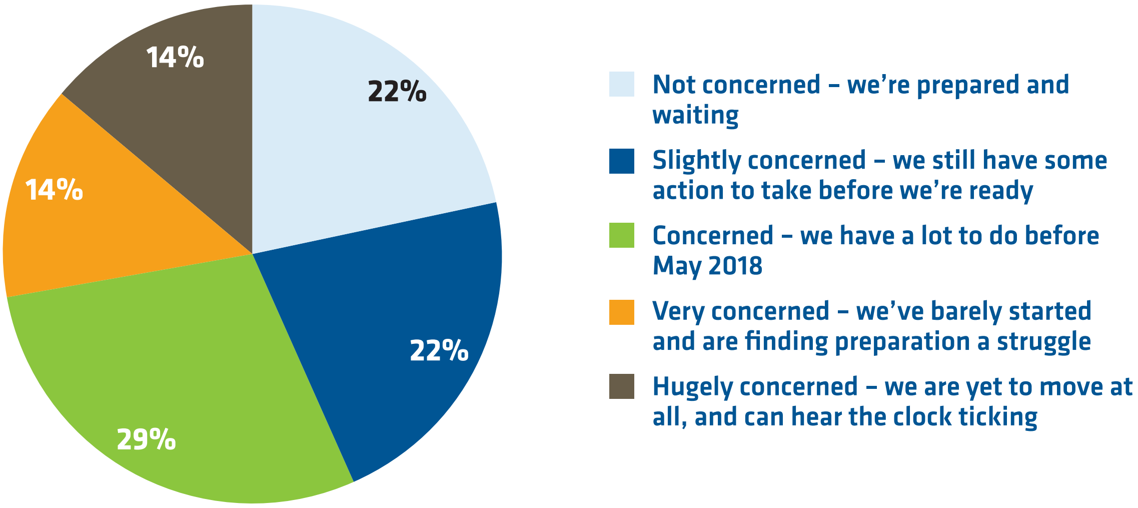
While many companies are enlisting the help of consultants to assist the transition, there are basic philosophies which a company itself can follow to begin navigating these new, data aware waters.

One approach is to look at GDPR as a form of agile development - not the waterfall approach it may now represent. Rather than linking it to your next project and in six months moving onto the next trend, GDPR needs to be embedded forever into an organisation, and to permeate everything you do in future.

That, of course, takes time. But with the right approach and the right combination of tools, solutions and policies, building a sustaining framework to support GDPR and other regulations is possible. The problem is not so much to do what is needed, but more on how to implement these policies, procedures and tools so they do not get in the way of new initiatives that the business needs to not just survive, but thrive in today's digital economy.

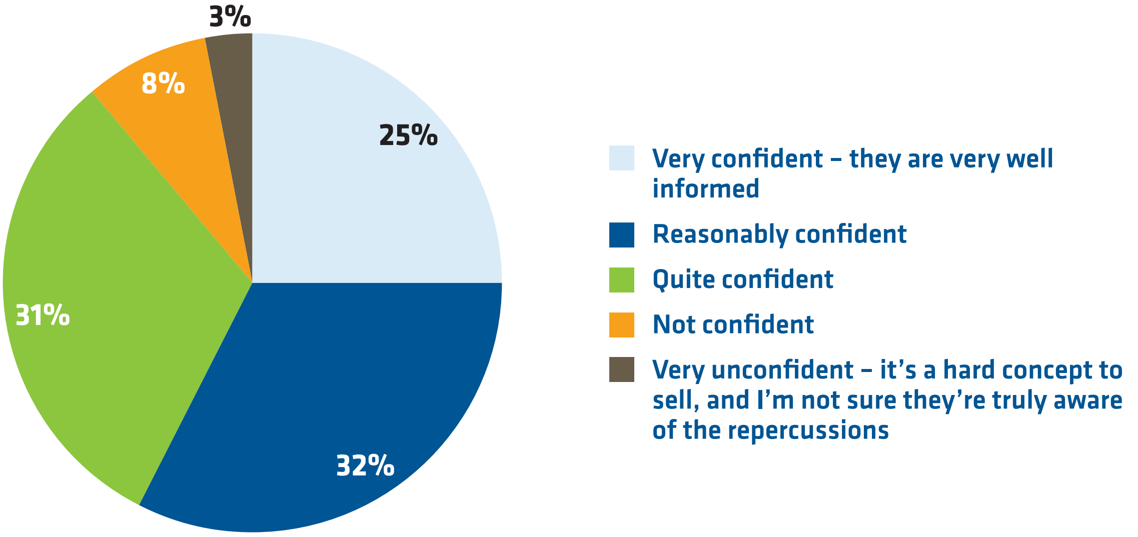
It's not a particularly positive start to the research, as Figure 1 shows. A total of 66 per cent of those questioned rate themselves between "concerned" and "hugely concerned" about GDPR, all happy to agree to statements that they've barely moved and have a huge amount to do in order to get their houses in order.

Fig. 1 : How concerned are you about GDPR?



At the same time, the 44 per cent of respondents "slightly" or "not" concerned seems, perhaps, an overly optimistic response. This becomes especially apparent when we begin to dig down into some of the realities of specific preparedness.

Fig. 2 : How confident are you that the board of your organisation are aware enough of GDPR to be ready to act?



The response when it comes to discussing board readiness for GDPR is a little more encouraging, perhaps, as Figure 2 shows over 57 per cent of respondents boasting "very" or "reasonable" levels of confidence that the higher-ups in their organisation are ready to act. Of course, there's a critical difference between a board that feels ready and an IT department that actually ***is*** ready, but nevertheless it's encouraging that so many IT decision-makers feel that their board-level GDPR dialogue is up to scratch.

Keeping storage GDPR-friendly

Of those surveyed, 72 per cent keep their data in mainframes, suggesting this tried-and-tested approach is still working for the vast majority of the enterprise. However, GDPR needs to know more than this, and it will be necessary to carry out a full audit of where data is stored in order to truly know. Respondents were also surveyed to find out in detail where they tend to store their data.

Fig. 3 : Which of the following types of facilities do you use to store and process data? (Select all that apply)

Data centre or server room based in Europe	68%
Private cloud	56%
Third party co-location / managed services facilities	40%
Public cloud	57%
Data centre or server room based outside Europe (including mirror and replication facilities)	36%
Hybrid cloud	23%
Third party contact centre	23%

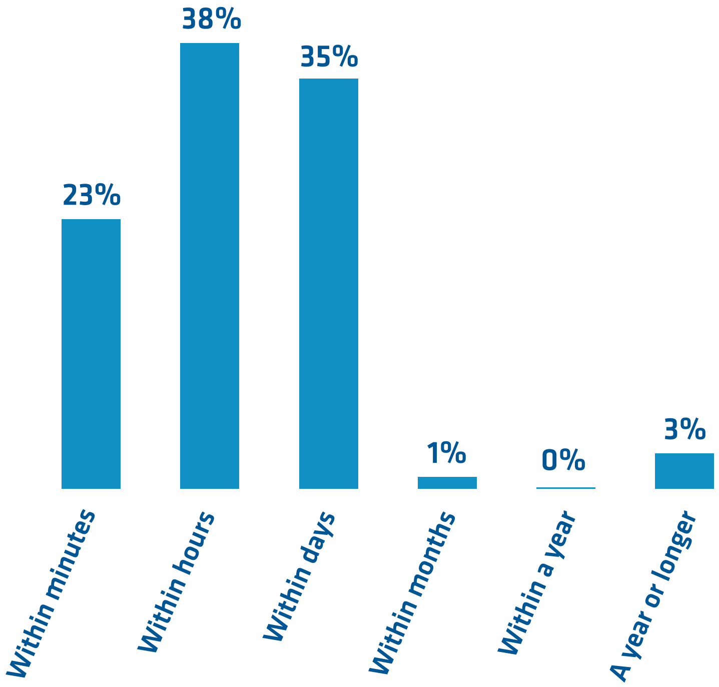
Identity and access management

The survey (Figure 3) showed that 46 per cent of respondents use single sign-on only, with 51 per cent using multiple sign-ons. A fringe three per cent used a mix of both depending on specific systems and use cases.

GDPR won't be letting security breaches sit unnoticed quite so easily, and with so many systems in play, many of which – such as Salesforce – exist in the cloud, or other applications and solutions hosted in enterprise clouds such as AWS and Azure, keeping tabs on identity is absolutely paramount.

After all, security breaches will most often be carried out by finding a user with system admin privileges and then using their identity to execute a change. Many think they have this issue covered, but recent breaches in public sector departments such as the NHS or high-profile private sector victims such as Sony and TalkTalk show this is very often not the case.

Fig. 4 : On average, when an employee leaves your organisation, how quickly are they denied access to systems?



Over a third of respondents admitting it takes them several whole days to remove user credentials of organisation leavers (Figure 4) doesn't help this situation. The almost three percent who admit to leaving these credentials lying around for a year or longer before deleting them is really quite shocking.

At the same time, when asked to state their level of confidence that they'll be able to locate and delete all the personal data on an individual within a strict time limit by the time GDPR becomes law in May 2018, respondents identified an average score of 3.6 along a slider bar where 1 was "not confident" and 5 was "very confident".

Luckily, software now exists to track leavers and remove them automatically, not to mention giving business management or HR direct access to systems to sweep details themselves, as opposed to leaving it in the hands of IT.

Tied up with single sign-on solutions, clean-up for those leaving organisations can become a much simpler – and more automated – process indeed.

Personally Identifiable Data – what it is and how to safely store it?

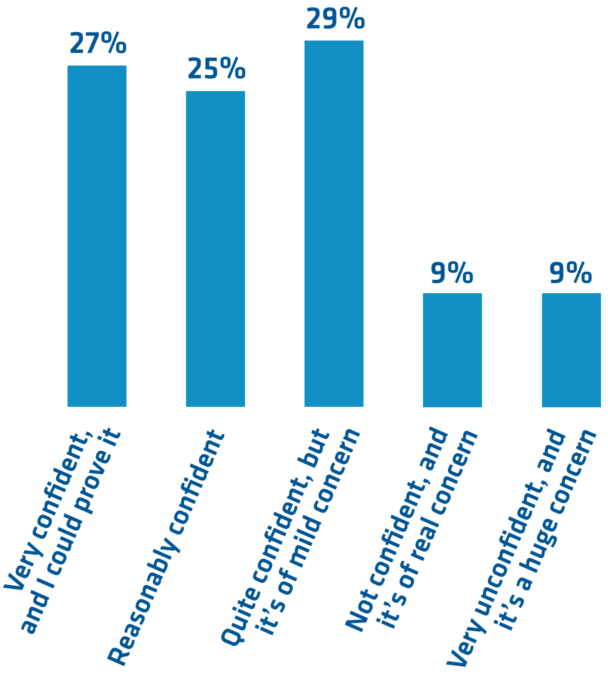
GDPR significantly ramps up the legal data privacy rights of customers, not to mention the requirements placed on organisations that retain, sell or share customer identities. And that applies to every single company doing business with customers who are citizens of the European Union.

The definition of PII – personally identifiable data – differs from simple "personal data". It's a term that originates in the US, and refers specifically, as the name suggests, to a range of data that is sure to identify a person closely. Name, address, date of birth, credit card and other financial information and social security (or in the UK, National Insurance) numbers are all included here.

GDPR's meaning of personal data, on the other hand, has a much wider purview, and is information more along the lines of social media posts, photographs, transaction histories and IP addresses – as well as everything just listed in PII.

It's clear that correctly storing PII is of paramount importance.

Fig. 5 : Are you confident that sensitive data and PII is stored in places where only you/your organisation can access it?



Luckily, Figure 5 shows that over 50 per cent of respondents are "very" or "reasonably" confident their PII is secure. But having said that, what we've learned so far suggests that confidence and the reality of the situation may involve something of a gulf.

Solutions now exist to take PII data and store it completely outside the applications and hardware upon which it's being processed – even in the cloud – and thus an extra layer of security is added, with the extra effect that all the PII data is stored in a completely separate system.

Software development and GDPR

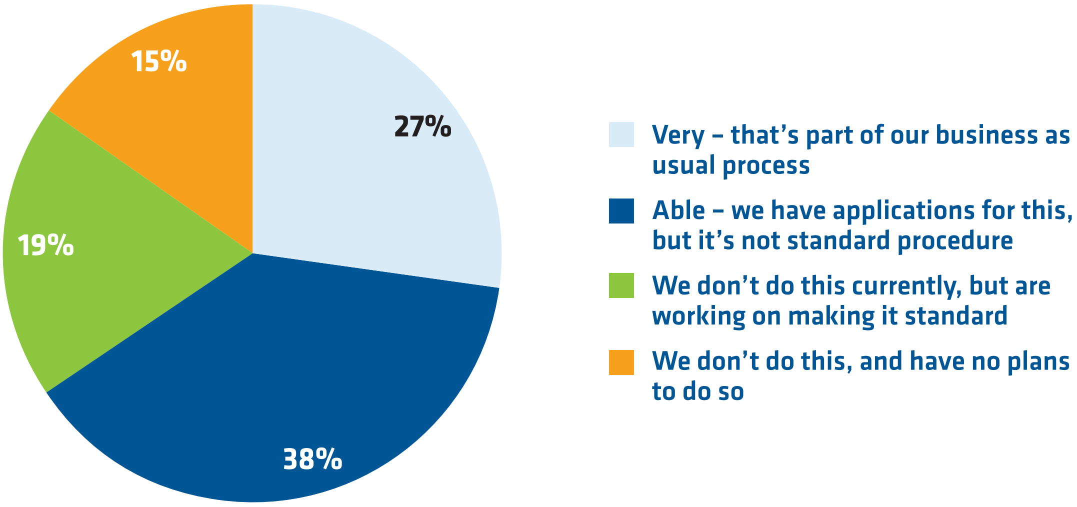
Survey respondents asked to grade the security of their API environment out of 5, 1 being "insecure" and 5 being "highly secure" reported an average of 2.7/5. This arguably doesn't feel robust enough for an increasingly API-based model of data access with GDPR as a backdrop.

GDPR will require organisations to explore new channels and methods for communicating with data sources, in order to support fundamental rights such as data portability per customer, including the easy transference of data to third parties, all via multiple file systems and sources.

It's going to be hard work, but again tools already exist to manage and control data flow through data processors – even through cloud integrated third party software – and help put minds at rest.

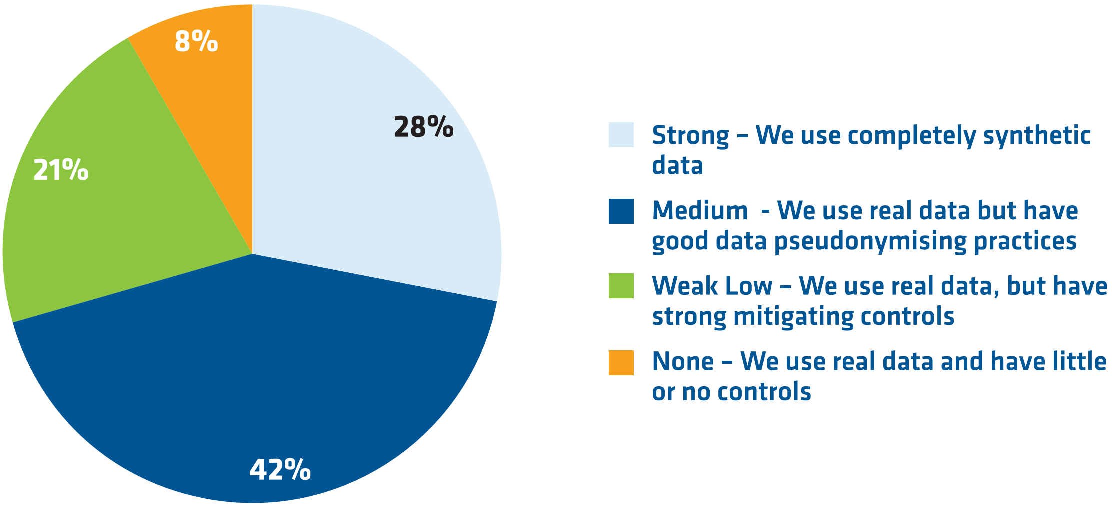
Looking into such solutions is recommended, particularly as the Internet of Things introduces a vastly growing environment of API-led data management.

Fig. 6 : How able are you to detect PII (Personally Identifiable Information)/sensitive data-vulnerable code during development?



As shown in Figure 6, well over 60 per cent of respondents believe they're in a good position to detect PII data-vulnerable code during software development, but how are they achieving this? Respondents were next asked about data leakage protection during the software testing stage, see Figure 7, *next page*.

Fig. 7 : How focused is your data leakage protection in software testing?



When discussing the results shown in Figure 7, it's important to identity that synthetic – or anonymised – data is not the same as pseudonymised data. The basic difference is that pseudonymisation takes the most identifying fields within a database and replaces them with artificial identifiers. For example, a name may be replaced with a unique, possibly randomised, number. Sometimes by design, however, pseudonymised data can be tracked back to its original form.

Anonymisation, on the other hand, is defined by the ICO as "the process of turning data in to a form which does not identify individuals and where identification is not likely to take place".

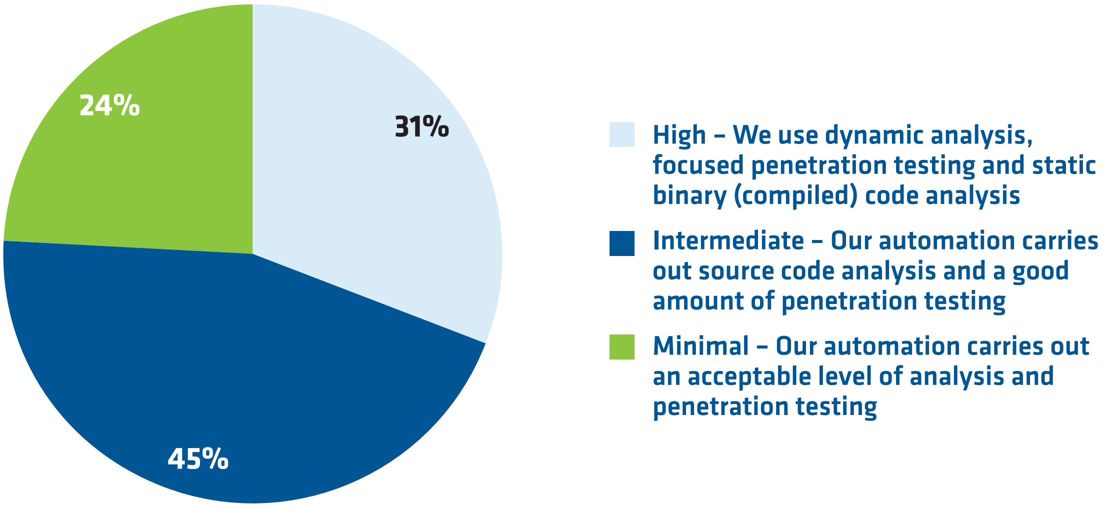
While this may sound similar, the difference is that anonymised data is non-human readable, and irreversible. One-way (preimage resistant) hashing is just one example of the permanency that can take place.

In most forms of development environment, it can be argued that there is little excuse for not using anonymised data, but this again requires picking the correct tools to generate the synthetic, anonymised data to begin with.

Synthetic data is surely a more painless approach in the long run, rather than the ever-growing risk of pseudonymised data missing the mark with GDPR if it happened to be reverse-engineered back into useful data by malicious actors. It's also worth remembering that pseudonymised data is still required to be treated as personal data, whereas anonymised data – if anonymised correctly – is not.

With less than 30 per cent of respondents currently taking the anonymised route and another 21 per cent still using real data with "controls", there is definitely work to be done.

Fig. 8 : To what level does your DevOps automation include software vulnerability testing?



Asked about DevOps automation and vulnerability testing (Figure 8), almost half of respondents revealed that focused penetration probably isn't carried out enough, and static binary – i.e. compiled – code analysis may not take place.

It's often common to carry out binary testing once applications are already live, or else leave vulnerability tests to extra, external software such as firewalls, which can massively slow down a user experience.

Choosing a software solution that tests binaries isn't just an essential notion for GDPR, it's also good practice for software development generally as DevOps, and agile development generally, becomes a more widely-used approach.

Conclusions

If the notion of GDPR approach as an agile methodology was the thesis for this paper, the conclusion surely has to be that there is much work still to be done by the enterprise. From an obvious mismatch between the belief of board preparedness to actual preparedness, drilling down into the specific areas of storage, security and development, there's a feeling that while organisations are now aware of the risks and the necessary philosophical practice that has – or will – become action, there is still an absence of that 'organic' inclusion that should make GDPR part of the furniture.

While those caught out will be able (or ordered) to appoint a Data Protection Officer (DPO), as well as served the usual enforcement notices and, subsequently, fines if improvement is not delivered, it's quite clear that a greater ground-level knowledge of aspects such as anonymization of data and options around PII storage and processing could negate this need by preventing problems in the first place.

"If you're genuinely acting in the best interest of the customer, then GDPR is almost a side issue because by default you are already doing that," Sainsbury's CDO Andy Day told *Computing* recently. There really seems no better message – build GDPR into your organisation's thought processes, toolsets and software solutions, and your risk will begin decreasing instantly.

But in order to choose the right tools and software, look at GDPR as a selection of individual moments –ask questions. ***How do I best want to secure my data, and which kind of data? Which milestones along the development journey are most at risk?***

By finding the specific answers to these questions, you'll also begin to ask the right questions about how to solve them, and with that, begin ticking the GDPR box for each and every task at hand.

About CA Technologies

CA Technologies helps customers succeed in a future where every business—from apparel to energy—is being rewritten by software. From planning to development to management to security, at CA we create software that fuels transformation for companies in the application economy. With CA software at the center of their IT strategy, organizations can leverage the technology that changes the way we live—from the data center to the mobile device. Our software and solutions help our customers thrive in the new application economy by delivering the means to deploy monitor and secure their applications and infrastructure.

For more information:

Visit: www.ca.com