

Data privacy: What every manager needs to know

July 2018

As companies begin to follow the principles of GDPR, the European Union directive on data protection, they must better understand what personal information really is and how to properly manage it.

In this episode of the *McKinsey Podcast*, McKinsey partner Kayvaun Rowshankish and associate partner Alexis Trittipò speak with Simon London about the European Union's newly implemented General Data Protection Regulation (GDPR), what it means, who it affects, and how companies can better manage personal data.

Podcast transcript

Hello, and welcome to this edition of the *McKinsey Podcast*, with me, Simon London. Whatever business you're in, like it or not, you're in the data-privacy business. The data you collect on customers, employees, prospects, even visitors to your buildings or your websites is increasingly subject to rules and regulations. The new European data-protection directive, GDPR, is part of the story but it's not the whole story by any means. Joining me today to discuss the issues are two McKinsey consultants who spend their time working with clients on exactly these issues related to data, analytics, technology, privacy, and risk. They are Kayvaun Rowshankish, who is a partner based in New York, and Alexis Trittipò, an associate partner, also based in New York. Kayvaun and Alexis, thanks very much for joining.

Alexis Trittipò: Absolutely. Excited to join you.

Kayvaun Rowshankish: Pleasure to be here.

Simon London: Data privacy: It feels like it's becoming a general management topic. Data in general, but data privacy as part of that. This definitely feels like something that, if you are a general manager, certainly in marketing, in sales and technology, but probably a whole bunch of other areas too, you have to know about this.

Alexis Trittipo: As we look at organizations, across industries, personal data is managed at all levels of the organization. It's not just a problem for the chief data officer or the chief information security office anymore.

This is a problem for HR. It's a problem for your customer-service representatives. It really spans the organization in terms of who touches personal data and who will need to, quite honestly, understand the regulations that are coming around.

Kayvaun Rowshankish: Every institution is heading in a direction where they are data-centric institutions. If you just take Amazon as an example. Sure, they stick physical products in the mail. But 90 percent of running their business is around data and data of individuals and data that represent the products that they manage. That's true of most sectors now. The other thing that's raising this to the management level is the fact that the scope of PII [personally identifiable information] is quite ambiguous. We might have thought of it previously as just a customer's name, maybe their social security number, other information about their whereabouts.

But now, because there's so much sort of behavioral data that's being captured, there are transactions and so on that are being captured, and multiple ways using advanced analytics to derive that information, it's just become a much more complex and opaque space.

Simon London: One of the big talking points at the moment is that the European data-protection regulation, GDPR, went into effect at the end of May. I suspect that a lot of managers are only now getting to grips with the operational implications of that. There are probably a lot of institutions, frankly, still scrambling to get compliant. So, Kayvaun, could you give us a quick overview of the contours of GDPR?

Kayvaun Rowshankish: In simple terms, the directive aims to protect individuals. And by "individuals," the focus is European individuals. But that doesn't mean the institutions out of Europe are exempt from this. European individuals travel, they have relationships with institutions outside of Europe, and there are third-party relationships with institutions that handle European individuals' data that then bring the next layer of institutions into scope as well. There's a pretty broad set of firms that are affected.

What the directive aims to do, again, in fairly simple terms, is it forces these institutions to put structure and discipline around what PII, personally identifiable information, actually is, where it is, put more choice in the hands of the individuals who the PII relates to, so that they can decide if they even want you to have it, or how to manage it (see sidebar, "The GDPR: Key facts").

THE GDPR: KEY FACTS

The scope of the European Union General Data Protection Regulation (GDPR) is broad, covering personal information that can be linked to an identifiable individual (such as national identification number, employee authentication, payment-transaction history, and date of birth) in any format (structured or unstructured) and in any medium (online, offline, or backup storage).

The regulation is designed to protect the privacy of EU residents by introducing stringent consent requirements, data-subject rights, and obligations on organizations that gather, control, and process data. Its core requirements cover the following:

Record of activities. Organizations should maintain a record of data-processing activities and be ready to present it to the regulator at any time.

Legal basis for data. All data processing should have a legal basis, such as the consent of the data subject or the need to fulfill a regulatory or legitimate business purpose.

Rights of data subjects. Data subjects are imbued with rights that organizations must honor such as the right to be forgotten (or, more accurately, to data erasure), the right to data portability, the right to object, the right to revoke consent, and the right to restrict processing.

Security. Organizations should protect data through a set of controls, such as encryption or “pseudonymization,” and have effective operational procedures and policies for handling data safely.

Third-party management. Vendors and suppliers, including outsourcing partners, should be required to protect personal data and should be monitored to ensure that they do so.

Privacy by design. Data protection should be included in the business-as-usual processes such as with any organization planning a new technology, product, or service from the beginning of the development process.

Breach notification. Data breaches likely to result in high risk to individuals' rights and freedoms should be reported to the authorities within 72 hours and subsequently to the data subjects as well in certain cases.

The new regulation is enforced via national supervisory authorities within the European Union that are granted wide-ranging enforcement powers and sanctions, such as the power to ban data processing. The fines for failure to comply are high, as much as 4 percent of annual worldwide revenues. The GDPR also allows individuals to seek civil actions (including class-action lawsuits) against organizations that violate their data-protection rights.

While GDPR is the most expansive regulation of its kind to date, there has also been movement in other geographies to increase protections around personal data. In the United States, for example, the state of California recently passed the California Consumer Privacy Act of 2018, which holds organizations to similar standards and imbues data subjects with similar rights.

It forces those institutions to minimize how much of that data they actually store [and process], puts controls around the PII that they store and process, and then enforces a set of governance and processes around that data, which includes accountability models like putting a privacy officer in place as well as a set of processes to interact with regulators and individuals

either in terms of managing their rights or responding to breaches and other types of events in this space.

Then lastly, the other thing that this regulation does is it puts in place some fairly severe penalties if you get it wrong. Quite explicitly, there's a penalty of up to 4 percent of global revenues, so you can be fined anything up to 4 percent of global revenues if you get this wrong. But it also puts in place, because this is going to become a legal requirement, the opportunity for individuals to take civil action against those institutions if they get it wrong.

Alexis Trittipio: GDPR is the first large-scale regulation of this size that's really aimed at protecting individual rights around personal data. The control that it aims to give individuals over their data in things like, I can have it deleted, I can ask to see what a company has on me, at this scale, is unprecedented. It's a very interesting test case to see how corporations will react. We'll see over time, both how this plays out in Europe, but also how this plays out globally.

Simon London: Is it right to say that even though GDPR is from Europe, certainly for any large global organization, it becomes the de facto global standard?

Alexis Trittipio: It's hard to say global standard because what we're seeing right now as folks are looking to implement it is they're largely implementing it for their European-based operations and where they touch EU-based customers and clients.

But most organizations are thinking about, "If I have to have this level of controls, this level of protection, this level of processes around personal data for Europe, how do I think about that more broadly? And how do I expand that?" And I don't think organizations are there yet. But as we think about it, this will be, globally, the highest standard around controlling and keeping personal data safe.

Kayvaun Rowshankish: There's a fair bit of noise around this issue as to whether the US should adopt consistent standards and just implement them across the US or if the needs of the US market are sufficiently different that they come up with something different. I think the same is probably true of other regions.

We do expect that there will be other regulations emerging covering regions that GDPR doesn't cover today that aims to address similar issues. If you look at history and regulations, things that came out of Europe first, the US chose to do something different.

If the same thing happens here, it's going to cause all kinds of problems. Because the systems that global institutions use for managing data of European individuals are the same systems they use for managing US individuals' information. If you have to start applying different standards and controls to those systems and processes, it's just going to cause a lot of fragmentation and redundancy and inefficiency, potentially leakage, and distraction of mind share because you've got too many different standards to deal with trying to address the same thing.

I do think that it would be ideal if there was some consistency and this became somewhat of a default that others tried to embrace and stretch out to other regions. But, as I say, if the past is anything to go by, it does give me pause that that would happen.

Simon London: One of the things that I read about GDPR is that it is principles based [exhibit]. Looking at a couple of principles laid out, these are pretty broad. I wonder, could you just give us a couple of examples of what it really requires of companies?

Alexis Trittipio: As you think about a principles-based regulation like GDPR, it leaves a lot open to interpretation. The principles are quite high-level. Each individual company will have to figure out how they're interpreting that, what the scope is, and what that means for them. And we'll see over time as regulators react to that.

One is around storage limitation. This is a principle that says basically, "Don't keep data longer than you need it for an active reason that you've told the data subject." And when I say "data subject," it could be a customer, an employee, it's a person like you or me.

This is contrary to the ways companies manage data today. Typically, I think, "Oh, I want to collect the data. I'm going to keep it for as long as I want. I'm going to use it in a bunch of different analyses." I don't think about, unless there's a regulatory reason, deleting that data. I just keep it in my central data warehouse, or, more likely, in various warehouses.

Exhibit

The General Data Protection Regulation sets out guiding principles for data protection.

Principle	Explanation
Lawfulness	Data should be processed only when there is a lawful basis for such processing (eg, consent, contract, legal obligation)
Fairness	The organization processing the data should provide data subjects with sufficient information about the processing and the means to exercise their rights
Transparency	The information provided to data subjects should be in a concise and easy-to-understand format (eg, the purpose of consent should not be buried in a lengthy document of terms and conditions)
Purpose limitation	Personal data may be collected only for a specific, explicit, and legitimate purpose and should not be further processed
Data minimization	The processing of personal data should be adequate, relevant, and limited to what is necessary in relation to the purposes for which those data are used
Accuracy	Data should be accurate and kept up to date
Storage limitation	Data should not be held in a format that permits personal identification any longer than necessary
Security	Data should be processed in a manner that ensures security and protection against unlawful processing, accidental loss, damage, and destruction
Accountability	The data controller is responsible for demonstrating compliance

Source: Regulation (EU) 2016/679 of the Council of the European Union, European Commission, and European Parliament

Companies are really going to have to keep track of how long they've had data and actively go about a process of data deletion. One of the things that's very challenging is how do I reconcile where I have to keep something for a legal reason, or for a tax reason, or a compliance reason, and where I need to delete it for this principle of limiting the length of time of data storage. That's something that companies are going to have to grapple with.

Another example is limitation around the purpose. The regulation requires you to tell the data subject why you're collecting the data. And it has to be collected for a specified, legitimate purpose. Then once you've used it for that legitimate purpose, you can't really use it for anything else, according to the regulation. So if I've said, "Hey, I'm collecting this information so I can open a bank account for you," for example, I can't then use that information to do marketing or other things. As companies think about this, the types of disclosures they have to make to data subjects will become important so that they can use the data for business purposes that they'll need it for.

Kayvaun Rowshankish: The other thing that I think most institutions are grappling with is just defining scope. Just simply the starting point of, "OK, who are covered individuals? And how do I define what is PII related to those individuals?"

So, sure, customers are covered. Employees are covered. But as you get into sort of visitors to buildings where you have to enter some information to get access to the building, they may be leaving PII behind. You have noncustomers that are making inquiries. At what point do they actually become your responsibility to protect information that they're entering in your website, for example?

And what is PII? There's obviously name, address, social security number. But as you get into things like IP address or clickstream information, it's not clear that that would be covered necessarily, and it would be extremely hard to put the types of controls around this stuff that the regulation's asking for.

Simon London: Although GDPR is very principles-based, there are some very specific rights, aren't there? If you're an EU resident, you get granted some very specific rights under GDPR. Can you just talk a little bit about those?

Alexis Trittipio: That's right. And, again, when we think about the purpose of this regulation, a lot of it is to give folks more control over their personal data. And so those rights are things like the right to access. An EU resident would have the right to call up a company and say, "I want to see all of the personal data you have on me."

Another one is, the right to erasure, also called the right to be forgotten, which I can call up a company and say, "I want you to delete all the personal data you have on me." There's things like the right to portability. I can ask for my data to be transmitted to someone else.

There's the right to not be processed in a fully automated fashion, which says, "I want a person involved in the decision making on me. I don't want to have decisions made on me based on analytics or based on machines or based on robots."

All of those rights are very core to GDPR. As companies think through how to comply, being able to ensure that they can meet each of these data-subject rights, within the allotted time frame, which for many of them is 30 days to get back to the data subject, will be quite important.

Kayvaun Rowshankish: The other approach which is likely is that they will just determine some cost benefit of even having that person as a customer. And for many of those that say, “No, I don’t agree to you automating decisions around me. I don’t agree to you storing this type of information about me,” then many institutions will just say, “Fine. I’m sorry, I can’t have you as a customer.”

“In order to get GDPR right, in order to get privacy right, you need the entire organization to be moving in the same direction.”

Simon London: You mentioned the requirement [for certain companies] to appoint a chief data-protection officer as part of the new regulation. What is work the data-protection officers do? It sounds like quite a pivotal role in actually coming to answer a lot of these open questions.

Alexis Trittipio: The data-protection officer, or DPO, is the hub of the wheel when it comes to GDPR compliance, and they should be very much central when it comes to data privacy overall to help facilitate when data subjects come and say, “I want to exercise one of my rights.” To help facilitate when there’s a data breach and making sure that there’s the right notifications to regulators. To really ensure that there’s consistency of data standards in the policies and the procedures. This person is the accountable party for GDPR.

Kayvaun Rowshankish: What’s quite interesting about this role and has become quite a thorny issue is that the DPO is responsible to the agencies and the authorities, not to the institution that they’re employed by. Plus, there’s language in the regulations that say that they need to have very senior reporting lines, whether to the CEO or the board.

What it actually means, and I think what most are interpreting it to mean, is that they need access too. And to Alexis’s point, they are accountable. But in very legal terms. If there is a breach and there is a penalty and someone’s going to get in trouble, it’s the DPO typically that is most legally accountable first and foremost.

So it has created a lot of problems in the industry around figuring out where to put this individual in the organization, what responsibilities to appoint to them, where geographically they should sit, especially for non-European institutions that have a global presence. Should they be in the US? Or would it be more relevant for them to sit in a European legal entity, which might actually put them further away from the board? There’s all these kinds of complications.

Alexis Trittipio: We see it in different levels of the organization. Sometimes it's tucked under a CISO, a chief information security officer. Sometimes this person on peer with the C-suite, or C-minus one.

As companies think about it, they've taken different paths. And, again, this is one of those things where, over time, we'll see what works better and what doesn't. And I think we'll see some shifts to a more standardized way of people treating DPOs similarly.

The other thing that's important to think through is it's not just the DPO, the individual themselves, that drives this. In order to get GDPR right, in order to get privacy right, you need the entire organization to be moving in the same direction.

It's the DPO and the privacy team, but it's also legal, it's also your entire data organization. Those individuals, and having the right team, that's a collaborative team across the organization, to deal with this is the most important. It can't be one individual alone.

“We're seeing all kinds of capabilities of AI that can help here, whether it's, for example, around identifying PII and preparing it and cleansing it through an entity resolution.”

Simon London: This comes at a time when institutions of all stripes are pushing very hard to collect and process data, to take advantage of machine learning, artificial-intelligence [AI] technologies. There's a lot of activity going on just to be sort of running in the opposite direction, certainly to GDPR. It feels like quite bad news for companies and their efforts to really take advantage of data and find competitive advantage within data.

Kayvaun Rowshankish: That's absolutely right, certainly as we talk to data-science communities that have enjoyed the pleasure of pulling in mounds of data into big data environments to just explore. You're seeing them being particularly nervous about the constraints that are likely to impact their, let's say, creativity. They're going to have to put more control around the AI models that they're developing.

Alexis Trittipio: The easiest way to get around GDPR, if you still want to use the data, is anonymize and mask. If data is masked and anonymized, you can use that. You can use that and not have to have the purpose told to the data subject ahead of time.

Thinking about, “Where do I actually need to know who the individual is?” versus “Where do I just need to have anonymized data that I can use for my analyses?” That will be key in the shortcut to getting this right.

Kayvaun Rowshankish: Ironically here, if you take the kind of counterintuitive of not how is GDPR blocking AI, you can look at how AI is actually enabling GDPR. There's plenty of

opportunities for the two disciplines to come together in that direction. We're seeing all kinds of capabilities of AI that can help here, whether it's, for example, around identifying PII and preparing it and cleansing it through entity resolution and other types of capabilities that come out of AI, through auto cleansing or deletion or masking of PII.

The concept there is that you have various different interactions with clients that are now moving from, say, human sales teams or human call centers to being AI-driven interfaces like chat bots or automated call centers where, frankly, you're not in control of what customers are telling you. You may be capturing PII inadvertently because customers are just sharing, so the other factor that comes into place here is whether you can use AI to detect when they're sharing PII, and either mask it immediately or delete it, or find some other way of blocking it so that you're not putting the firm at compliance risk.

Simon London: I just want to push on this point around big data and machine learning. To a layperson, it almost sounds like creating a data lake by combining information about your customers from multiple sources and then running algorithms on that data with the intention of extending different offers to different customers. So Simon would get one type of offer, Alexis would get a different type of offer. It sounds to me like that could be a breach of GDPR.

Alexis Trittipio: I think you go back to the reason you collected data. You have to be explicit about why you're collecting data. As long as it's a legitimate business purpose. A legitimate business purpose can be providing the best and most appropriate products to our customers.

So if you've collected that data with the customer or the client understanding that, then I think that kind of analysis is allowed. Going back to what am I consenting to, as a data subject when you collect my data, and making those disclosures very clear that they're going to be used for the purposes of understanding and marketing products.

Kayvaun Rowshankish: If you just look at the extent you use a lot of these online applications, so many of them have been putting up these pop-ups as you access the site asking for your consent. I'm sure most people aren't actually reading much of that; they're just clicking, "OK." It's basically that that you're consenting to, that you can actually use that information to better tailor the sales process.

"I do think there needs to be a phase two and a phase three where you go into the backups and the paper files and actually fully extract that person."

Simon London: Which strikes me as being fine for data that's collected now or recently. The question is, is everything that's going into your data lake, was it collected under those types of conditions? Or is it historic data where the permissions in place at that time were not compliant with GDPR? You were not as explicit as maybe you could have been or should have been around the potential uses of the data.

Kayvaun Rowshankish: This is an interesting point because these consents that are being asked for, as I was describing, if they're written correctly, they would cover data that you've captured in the past. You may have them in backup, you may be using them for a variety of different purposes. And as you're asking individuals for consent to use their PII, most institutions are creating that consent statement in pretty broad terms. The issue is, what if they say no? And they actually say, "No, we don't want you to store PII, we want you to delete everything that you've got about us." I don't think I've met a single institution that has really figured out how to deal with that.


They're dealing with it through drawing false boundaries, let's say, around, "OK, anything that's in a live system, we will find a way of deleting that data about them. But if it's in backup, if it's on paper, then it's not as easily discoverable anyway, so we're not going to tackle that."

It's not completely clear that that is in compliance with how the regulators are thinking about it. It's still a bit of an ambiguous space.

Alexis Trittipio: And I think what we're seeing across industries, is, to your point, Kayvaun, this will have to be a phased process. Most companies say, "In our live systems, yes. We can delete all your information." I do think there needs to be a phase two and a phase three where you go into the backups and you go into the paper files and actually fully extract that person. It'll be interesting to see how organizations take that compliance from the more immediate "in my active systems today" across to the paper files from 20 years ago as well.

Kayvaun Rowshankish: And it's interesting, at the same time, there's a whole lot of digitization using OCR, optical character recognition, natural language processing, to extract the data that you want from PDFs and physical documents that it won't be long before you probably shouldn't be storing paper copies of things anyway.

Storage is getting so cheap that even the backups should be fairly accessible that if you wanted to perform against this rights to erasure, then it should be much easier than it is today in the coming years.

Simon London: So I think that's all we have time for today. Thank you very much to Alexis, and thank you to Kayvaun. And thank you to our audience for listening. To learn more about data privacy, GDPR, risk, and regulation, please visit us at [McKinsey.com](https://www.mckinsey.com). 

Kayvaun Rowshankish is a partner in McKinsey's New York office, where **Alexis Trittipio** is an associate partner; **Simon London** is a member of McKinsey Publishing and is based in the Silicon Valley office.