Article development led by **acmqueue**
queue.acm.org

Legal considerations and broader implications.

BY JATINDER SINGH, JENNIFER COBBE,
DO LE QUOC, AND ZAHRA TARKHANI

Enclaves in the Clouds

WITH ORGANIZATIONAL DATA practices coming under increasing scrutiny, demand is growing for mechanisms that can assist organizations in meeting their data-management obligations. In this context, trusted execution environments (TEEs), also known as *enclaves* or *secure enclaves* (the terms are used interchangeably here), are the subject of much interest.

TEEs provide hardware-based mechanisms (enclaves) with various security properties for assisting computation and data management. Broadly, TEEs are concerned with the confidentiality and integrity of data, code, and the corresponding computation. Because the main security properties come from hardware,

certain protections and guarantees can be offered even if the host privileged software stack (for example, the operating system or hypervisor) is curious, vulnerable, or compromised.³

The technology is growing in prominence. There are a range of hardware offerings, and the major cloud providers are beginning to feature enclaves in their services.

A Technology of Current Interest

TEEs have a rich history, based on long-established security principles. In the early 2000s the use of such technology gained some attention by way of “Trusted Computing” initiatives, which proponents argued were advantageous because they provided the means to enforce particular systems operations. These initiatives were met with controversy, given that the technology effected a loss of user control in the interests of vendors (see the “Application Contexts” section later in this article).⁴

Today the growing awareness of the likelihood and costs (financial, reputational, legal) of security and privacy breaches, along with the increased legal and regulatory attention to data and technology more generally, has resulted in more interest in TEEs’ potential to assist organizations in managing issues of data and its processing.

One key driver for this interest is the dominance of the cloud (that is, service-oriented computing in multiparty environments). This is because enclaves can provide a technical means for enabling isolation and assurance against the “untrusted” host (that is, the provider), other *tenants* (customers of cloud services), and other parties with which interactions occur. The industry has recognized this potential for enclaves to shield a tenant’s data and processing from others, including the cloud provider, by effectively removing the provider from the *trust chain*.^{16,22} Reflecting this, terms such as *confidential computing* are used to highlight cloud-related business cases for such technology.^{13,16} Enclaves can also work in a cloud context to assure

the integrity of systems, that certain processing occurred, that configuration and management policies are applied,⁸ and that billing accurately reflects the resources consumed.¹

Directly related to this interest in TEEs is the attention being paid to the privacy and security considerations of data analytics and ML (machine learning), which is increasingly occurring in the cloud. Work has been ongoing on enclaves for enabling secure multiparty computation on untrusted platforms,¹⁹ as well as for privacy in large-scale data

analytics (for example, SGX-PySpark¹²) and enabling more privacy-aware machine learning as a service.¹¹

This article focuses on TEEs for the cloud, because the cloud represents the dominant form of providing applications and data processing infrastructure.

Enclaves and the Law

The legal requirements and obligations of those designing, deploying, or using technical systems for processing personal data stem from a variety of sources. A prominent example is the

European Union's GDPR (General Data Protection Regulation), which is both high profile, and, though an EU regulation, also globally relevant—it not only operates across territories, but also serves as a model for the data protection regimes of many other jurisdictions, including the CCPA (California Consumer Privacy Act). Various aspects of GDPR have direct technical implications, including in relation to security, organizational accountability, and data protection by design.

GDPR has received much attention,

given the potentially significant penalties for noncompliance. Still, it is only one example—legal obligations with technical implications can flow from other sources, such as contractual agreements or product-liability regimes.

It follows that demand is growing for mechanisms that assist organizations in meeting their legal responsibilities. TEEs are seen as one technology to help with such concerns. Indeed, vendors have explicitly characterized the technology as enabling compliance with various regulations (see the “Cloud Computing” section later in this article). Despite these assertions, however, the extent to which enclaves assist with legal obligations requires exploration. So far, neither the implications of such protections in a service-provider context nor the potential for this technology to reinforce the dominance of the large providers have been given much consideration.

Accordingly, this article provides a cross-disciplinary analysis of how the functionality of TEEs relates to legal obligations in the context of the cloud. The article presents a high-level overview of the common security properties provided by enclaves and introduces some of the relevant legal landscape. The focus is on data protection (specifically the GDPR), as that represents a key area of law relating to data concerns.

The article then explores how enclaves can assist an organization’s legal obligations, responsibilities, liabilities, and compliance concerns in a service-provider context, considering the benefits for the parties involved. It shows

that although TEEs might assist in managing some legal obligations, they will not resolve issues of data protection compliance—despite what some marketing materials might imply.

Further, despite some benefits for tenants, the technology is shown to favor the interests of the service provider (similar to how previous instances of the technology were seen as protecting vendor interests), and the implications of this are considered. Finally, the technology’s potential as a foundation for enabling compliance and accountability regimes is considered.

Technology Overview

At a high level, a TEE (or *secure enclave* or just *enclave*) can be described as an execution environment that isolates the code and data inside it from the host’s software. The security guarantees are backed by hardware (for example, CPUs), protecting the confidentiality and integrity of the data and code within. Cryptography often plays a key role in the security guarantees.

Conceptually, Trusted Computing is not new. In the 1980s the U.S. Department of Defense Trusted Computer System Evaluation Criteria,²⁸ known as the “Orange Book”, provided the basic principles of trust for sensitive applications. This included a definition for a TCB (trusted computing base) and generally established the need for minimalism, compartmentalization, and the verification of systems-critical hardware/software modules. Efforts to develop trusted computing technology continue today, both in terms of defin-

ing new requirements and the practicalities of their implementation. Many hardware vendors and service providers now incorporate such technology into their offerings (see the “Application Contexts” section later in this article).

Security properties. TEEs reflect a broad set of security features and concerns, as outlined here. The accompanying figure presents the purported features of some prominent implementations.

In-enclave protection. TEEs can provide a shielded execution environment that protects (to various degrees) data and code against unauthorized access and modification by external software. This includes protections from the host operating system or hypervisor that otherwise would generally have the potential for full control of and visibility into the tenant’s code and data. Note that of those listed in the figure, only Intel SGX and AMD SEV support runtime memory encryption.

Remote attestation. This feature securely verifies the internal state of a platform and the in-enclave application. It gives certain guarantees about the enclave before the transfer of “secrets” (keys/certificates, data, or code) to a remote platform and enables verification that the code and data running in the enclave are correct—for example, not modified or tampered with. It can be achieved either statically at boot time or dynamically at runtime to establish a *dynamic root of trust*, which can serve as a foundation for providing various systems-level guarantees.

Sealing. Persisting an enclave’s secrets to a disk in a protected manner enables the retrieval and reinstantiation of in-enclave state, perhaps as a result of system reboot, power interrupt, or being in line with application requirements. Sealing requires binding the secrets to a specific enclave’s state or device.

Secure boot. A secure boot defines and ensures a chain of trust regarding the enclave images, operating system components, and configurations. Often this entails loading immutable and verified images to an enclave.

Isolated peripherals. Secure peripheral sharing (for example, NICs) and isolated I/O paths are enabled via fine-grained control over hardware resources or using cryptography-based mechanisms.

Overview of some prominent TEEs.

TEE	Target ISA	Security Features						
TPM/vTPM	Multiple Targets	●	●	●	●	—	—	—
Intel TXT	x86_64	●	●	●	●	—	—	—
Intel SGX	x86_64	●	●	●	—	—	—	—
ARM TrustZone	ARM	●	—	●	●	●	●	—
Sanctum/KeyStone/MultiZone	RISC-V	●	●	●	●	—	●	●
AMD SEV	AMD x86	●	●	●	—	—	—	—

● provides property
 ● supports partially
 — does not support

In-enclave protection
 Remote attestation
 Sealing
 Secure boot
 Isolated peripherals
 Side-channel resistance
 Open source hardware

Side-channel resistance. This feature mitigates information leakage through various software- or hardware-based side-channel attacks.

Technical considerations. Leveraging TEEs raises several practical considerations, some of which we outline below.

Usability. Properly leveraging the capabilities of enclaves requires expertise, including an understanding of the particular technology's role, features, and limitations in a specific context. TEEs operate only to directly protect that code and data residing within the enclave, although they are often used as a foundation to provide broader systems protections and guarantees. Developing enclave-assisted applications involves considering which aspects require and are most amenable to protection and compartmentalizing the application to use the TEE's functionality appropriately. Tools are being developed to support running legacy applications inside enclaves without modification (for example, SCONE⁵) or running applications on heterogeneous enclaves (Microsoft's OpenEnclave²⁰). One must also determine how best to manage and protect the code and data running outside the enclave.

A growing TCB. Traditionally, enclaves were designed to support very specific and small operations (for example, cryptography or key management). The ideal is for enclaves to contain an uncomplicated and verifiable in-enclave software stack, with minimal interactions with the outside environment, so as to offer strong security guarantees and assurances.²⁸ In practice, however, enclaves are used in much more complex application architectures such as web services, auditing, and privacy-aware data analytics.

The risks associated with service provision and software stack vulnerabilities, and the difficulties in engineering and refactoring applications into trusted and untrusted compartments to leverage enclaves, have led to enabling in-enclave support for a larger portion of an applications' dependencies, including entire language runtimes and operating systems.²⁷ For the cloud, a larger TCB allows more complex applications to be deployed by tenants, thereby making the cloud more attractive to them. Again, however, security guarantees and assurances are

generally significantly lowered by having larger and more complex TCBs.²⁹

Security vulnerabilities. TEE hardware provides only particular security properties—the software or system design must properly leverage these to benefit. Insecure or buggy in-enclave code will often not (depending on the issue) mitigate security risks and, indeed, could actually render a worse outcome as the enclave works to protect the problematic computation (for example, acting as an enclave that protects malware²³). As already mentioned, increasing the size and complexity of what resides in the enclave and various layers of interactions with the untrusted world causes a range of vulnerabilities that can jeopardize the enclave's security properties and guarantees.^{26,29} Further, there may be vulnerabilities regarding the TEE itself (architecture or supporting services)—various attacks have been demonstrated¹⁸—and, as a commodity technology, an enclave issue can result in insecurity at scale, as the Spectre/Meltdown processor vulnerabilities demonstrate (see www.meltdownattack.com).

Note that many enclave offerings are proprietary and closed (to varying degrees), meaning that it can be difficult to ascertain precisely what is offered and how they work, and to assess and validate their claims. Moves toward more open access to software and hardware stacks would assist such interrogations.

Standardization. The growing adoption of enclaves means they will increasingly form part of a wider technical ecosystem. Therefore, standardization is important for facilitating adoption, consistency in functionality and use, as well as allowing interoperability and management in increasingly heterogeneous environments. Various standardization initiatives are under way, including by GlobalPlatform, the Trusted Computing Group, and the Confidential Computing Consortium, with many different areas warranting consideration.

Application Contexts

As a general-purpose security technology, there are many potential applications for TEEs. In the early 2000s came the various “Trusted Computing” initiatives, which aimed to enforce particular behavior and functionality on

a machine. They were controversial, seen as technologies serving vendor (rather than user) interests.^{4,21} Concerns included power asymmetries and anti-competitiveness, given that trusted computing effectively enabled vendors to control what users could and could not do on their machines, thereby facilitating DRM (digital rights management), the reinforcement of vendor ecosystems (lock-in), and so on.⁴ Note that this was in an era when software and processing predominantly occurred locally, on user hardware; however, as we explore, the concerns about who the technology benefits, as well as its wider implications, remain.

Today a prominent use of enclaves is providing security guarantees and assurances regarding data and its processing, and how they relate to broader computational infrastructures. The focus here is on the use of TEEs in the context of cloud computing, where various technical capabilities are offered as a service in a multiparty environment. Note, however, that enclaves are also increasingly embedded within a wide range of systems to underpin security functionality,⁹ and will likely play a role in supporting edge and IoT (Internet of Things) ecosystems.²⁴

Cloud computing. TEEs are increasingly marketed as a means of helping mitigate some of the risks associated with the cloud. Cloud computing entails the provision of computing functionality, resources, and infrastructure, delivered as a service.¹⁷ The cloud encompasses several parties: the cloud service *provider*, which offers the particular service(s) to *tenants*, who are the customers and direct consumers of the services.

A vast range of cloud services are available; in general, tenants use the cloud to provide or manage (aspects of) their technical infrastructure, including, the storage and processing of data. That is, cloud providers offer the infrastructure and components for supporting, hosting, and running customer applications, services, data analytics/ML, and so forth. Cloud providers leverage economies of scale, whereby the services and resources they provide may be shared among tenants. Often the cloud uses a pay-per-use model that is attractive for tenants, given that they can access compute resources on demand, as

needed, typically at a substantially lower cost and with lower overhead than maintaining such services in-house.

Security is a key concern in the cloud. Enclaves can allow a tenant's code, data, and processing to be better isolated and protected: from other tenants on the shared infrastructure; and, importantly, and from the cloud provider itself—though a provider might host particular data or computation, enclaves can limit a provider's access to such. This is important, particularly where the data, code, computation, or analytics might be personal or otherwise sensitive. Further, TEEs can also be used to give guarantees that certain processing occurred, certain management policies were enforced, and the like.

Note that the service provider integrates the enclave technology into its service infrastructure. In this context, TEEs provide certain controls and guarantees over the providers' systems, as opposed to the situations where a TEE operates as a control mechanism over the user's hardware/system.

Enclaves as a service. TEEs can be offered by service providers at different levels of technical abstraction, depending on the features of the particular TEEs employed and the particular use cases. In practice, enclave offerings tend to take the following forms: Trusted/trustworthy virtual machines (for example, Intel's Trust Domain Extension (TDX) or AMD's SEV (Secure Encrypted Virtualization)); an in-enclave LibOS (library operating system, for example, Graphene-SGX²⁷) or container (for example, SCONE⁵); enclave-assisted applications and analytics frameworks (for example, SGX-PySpark,¹² and VC3 [Verifiable Confidential Cloud Computing]²²); or a limited SDK and programming framework (for example, Google's Asylo and Microsoft's Open Enclave).

Enclaves can underpin a number of services in a cloud context. First, they can be used to support and protect specific operations—for example, for cloud-based key or credential storage, management, and verification. Alternatively, the TEEs of the cloud provider can be used to support custom applications, as developed by the tenants themselves, or more generally where a whole application, container, or operating system might be enclave protected (see the earlier section on “Technical

Considerations”). Allowing larger applications as well as data processing or analytics operations to run within an enclave may be attractive to prospective tenants. These offerings can support a range of tenant requirements, while serving to make some of the technology's functionality more widely accessible, thus recognizing that leveraging enclaves can require some expertise.

The major cloud providers already offer enclave-backed infrastructure services. Indeed, providers see the emergence of TEEs as facilitating the growth of cloud computing. With industry materials describing the technology as enabling “confidential computing” and removing the provider from the “trust loop,”^{13,16,25} enclave-backed offerings attempt to encourage otherwise risk-averse customers (tenants) to make more use of cloud services because of the security guarantees and assurances that the technology can provide.

In particular, the technology can enable computation to happen on provider infrastructure, without that provider having plaintext access to the code, data, and processing. This can be attractive where the nature of the data or its processing is highly confidential, personal, or subject to other constraints. The services are also marketed as providing integrity guarantees (that is, assurance the data has not been unduly modified, that the correct processing actually occurred, that machines were configured correctly, and so forth).

TEE offerings are also promoted to tenants as supporting the management of legal obligations. For example, Intel's description of IBM's cloud explicitly describes the technology as assisting with GDPR compliance.²⁵

In summary, enclaved-backed service offerings provide incentives for customers to use cloud services. Prospective tenants can leverage the cloud providers' significant resources and technical capacity to take advantage of the security and possible compliance benefits of the technology (and related services), while avoiding the potentially significant overheads of in-house infrastructure management.

Legal Context

The design, deployment, and use of systems, and the management and processing of data, are the subjects of

growing legal and regulatory attention. Various legal obligations and requirements—arising from data protection, security, and product safety legislation, as well as from contractual obligations and other liability concerns—are increasingly relevant for technical systems and will encourage technological responses. The adoption of technical measures (including for security) to meet legal obligations is often driven by the fact that failure to do so can result in penalties and restrictions on operations, among other repercussions.

The focus here is on data protection law, given its particular prominence and general relevance to systems and data concerns. Note, however, that other laws can be relevant and can also impose legal obligations and responsibilities regarding security and data management (see “Law as a Technology Driver”).

General Data Protection Regulation. This article's discussion is in the context of the EU's GDPR, given its profile and breadth. Though the GDPR is EU law, it has worldwide relevance; it can apply to entities outside the EU that process data on EU individuals (GDPR Art.3), and it's becoming a de facto worldwide standard, with other jurisdictions establishing laws along similar lines.

GDPR's framework. GDPR applies to the processing of *personal data*—that is, any data relating to an individual who can be directly or indirectly identified (GDPR Art.4(1)). This includes not only names or IDs, but also location data, device identifiers, online identifiers, IP addresses, and so on. Similarly, personal data can also exist where an individual may be identified from some combination of factors specific to the physical, physiological, genetic, mental, economic, cultural, or social identity of that individual. Important for this discussion is that encrypted personal data is often still considered personal data and is thus subject to GDPR (even if the entity holding the data does not hold the decryption keys or can't otherwise access the plaintext).⁷ This is because anyone with access to the plaintext could use it directly or indirectly to identify an individual.

As defined by GDPR, *processing* includes any operation or set of operations performed on personal data,

including, for example, collection, storage, adaptation or alteration, retrieval, and use (GDPR Art.4(2)). Note that this is broader than the standard computer science use of the term, as it entails far more than just computation.

GDPR establishes that those involved in processing personal data act as data *controllers* or data *processors*. Controllers are those entities that determine the means and purposes of processing personal data (GDPR Art.4(7)). Processors, on the other hand, are entities that process personal data on behalf of and under the instruction of a controller (GDPR Art.4(8)). In simple terms, controllers generally decide why and how the processing happens, while processors carry out particular processing operations for the controllers.

Whether one is a controller or processor is important, as it has a direct bearing on legal obligations and responsibilities. In many instances a company providing an application is the data controller, while others providing various technical services are usually data processors (that is, regarding the cloud, for example, often the cloud provider is the controller, though this will depend on the specific circumstances¹⁷). The role taken by a particular entity, however, is determined on the basis of who is actually doing what, regardless of any contract or other agreement between the entities.

Security-related obligations. GDPR places positive obligations on controllers and processors in relation to security and data management. GDPR mandates the principles of data protection by design and by default (GDPR Art.25), whereby technologies, applications, and processes need to be designed from the earliest stages with data protection in mind. GDPR also requires that both data controllers and processors take “technical and organizational measures” to ensure a level of security appropriate to the risks incurred in processing personal data (GDPR Art.32; Recital 83). These risks may relate to, among other things, loss, alteration, or unauthorized data disclosure or access.

Enforcement and penalties. The penalties for GDPR noncompliance can be significant. Supervisory authorities (regulators) are empowered to take various actions against both controllers and processors for failures to comply

GDPR places positive obligations on controllers and processors in relation to security and data management.

with GDPR’s requirements, including banning them from processing data (GDPR Art.58(2)), which could prove fatal to an organization. Both controllers and processors are liable for administrative fines of up to the greater of €10M or 2% of global turnover for certain breaches, including the failure to meet GDPR’s security requirements, or €20M or 4% of turnover for some classes of other GDPR breaches (GDPR Art.83(4)). In this way, GDPR seeks to be effective against even the largest companies.

Law as a technology driver. The GDPR is an example showing how law can introduce legal requirements and obligations that impact system design, and where compliance has technical implications. Though data protection law is particularly prominent and broadly applies to personal data processing concerns, other laws can also raise security- and data-related considerations. (For reasons of space these are not explored here.) Various parties in a technical ecosystem will face legal requirements and obligations, whether through contract, liability, statute, or regulation.

With the prospect of potentially significant penalties for noncompliance, whether stemming from data protection requirements or elsewhere, there are strong incentives to implement risk-mitigation measures beginning from the earliest stages of system design. Therefore, as a security technology offering a broad range of capabilities, TEEs have received considerable attention as a *possible* means for helping systems designers and operators meet legal obligations and requirements.

Legal Relevance and Discussion

TEEs might assist in meeting legal obligations primarily by: practically reducing risks relating to security (and data governance) or providing evidence (technical or otherwise) of steps undertaken to mitigate those risks, or demonstrating that the appropriate actions were taken.

In this way, TEEs can help meet some of the requirements set out in GDPR, as well as those from other regulations. Enclaves might also assist in meeting the security obligations that can arise from contractual relationships. Even where security requirements are not explicit

in law, security-oriented requirements may arise from best practices, adherence to industry standards, and so on. TEEs can assist with meeting these requirements in much the same way as they can assist with meeting similar obligations established by law.

The following sections consider in more detail how TEEs relate to obligations under GDPR in a cloud context. Note that some of the discussion may be applicable more widely.

Security-related requirements. Recall that GDPR requires data controllers and processors to take the appropriate “technical and organizational measures” to ensure a level of security appropriate for the risks of data processing (see “Security-related Obligations”). In determining the measures to employ, parties should take into account the current technological state of the art; the scope, context, nature, and purpose of the data and its processing and the associated risks; and issues around loss, alteration, unauthorized disclosure and access, among others (GDPR Art.32; Recital 83). GDPR explicitly envisages that appropriate measures might include encrypting personal data and mechanisms to ensure the confidentiality and integrity of processing (GDPR Art.32; Recital 83).

Because of their significant resources, it is argued that cloud providers are often better placed to deal with security concerns than companies managing their infrastructure in-house. Moreover, GDPR requires that processors can be used only if they provide sufficient guarantees that they have implemented measures ensuring GDPR-compliant processing, including in relation to security obligations (GDPR Art.28). In this way, enclaves, beyond potentially further enhancing a provider’s security posture, could provide an additional mechanism to help cloud providers offer such guarantees.

Conversely, a vulnerability in the cloud service (for example, an incorrectly configured enclave-backed service) potentially affects everyone using that service. There is, therefore, a significant reliance on cloud providers to implement TEEs correctly (as with other aspects of their infrastructure)—where the discovery and consequences of a vulnerability can be more widespread, affecting swaths of tenants.

Using TEEs, which include cryptographic means for isolating data and compute, does not remove one’s obligations under GDPR.

Ultimately, enclaves are but one measure that can help meet security obligations. By adopting technical measures such as TEEs—which can provide isolated execution, protected memory, remote attestation, and enabling other relevant functionality such as secure recordkeeping/logging (see “Enclaves: A Basis for Compliance and Accountability?”)—controllers and processors may be in a better position to meet their security obligations and mitigating or reducing their risk of penalty where breaches occur. Enclaves will provide only so much on their own, however; appropriate compliance regimes entail a holistic approach encompassing both technical and non-technical organizational measures.

Encryption and personal data. Encrypting personal data does not necessarily mean it is no longer personal data.⁷ Again, this is because the data will still relate to an identifiable individual and can be decrypted, either with the appropriate key or because of some vulnerability in the encryption mechanism—both of which could be discovered or compromised at a later date. Moreover, both the act of encrypting and subsequent operations performed with that encrypted data (including storage, retrieval, transfer, decryption, or use) will also be considered processing and subject to GDPR.

In other words, although encryption can mitigate some risks regarding the disclosure of personal data, it often does not affect the status of that data as personal (GDPR Recital 28). Organizations will therefore still have the same obligations for appropriately managing encrypted personal data as they do for personal data in general.

In this way, using TEEs, which include cryptographic means for isolating data and compute, does not remove one’s obligations under GDPR. This raises implications for those processing data on behalf of others, such as cloud providers. Those providing such services still have responsibilities under GDPR for encrypted personal data, even if they do not have access to the legible (plaintext) data or decryption keys, and importantly, even if they are unaware that the data is personal in the first place.^{7,17}

Analytics and machine learning. Data analytics and machine learning are areas of considerable commercial inter-

est. Note that despite the discussions concerning privacy-preserving analytics, the position under GDPR remains the same for analytics and ML as it does for any other form of processing. That is, analytics and machine learning are just particular forms of the wider category of operations performed on personal data that are collectively termed *processing*, and therefore it has the same obligations.

The cloud is described as having a key role to play in providing the infrastructure for enabling data storage, analytics, model building, among others. Though using the cloud will not relieve an organization's data protection obligations, there are responsibility-related incentives for prospective tenants leveraging enclave-backed processing environments for analytics and ML. As noted, controllers are obliged to use only those providers that can provide sufficient guarantees that their undertakings are in accordance with GDPR requirements—where, again, TEEs might assist providers in demonstrating their compliance. Further, using enclaves for cloud-based analytics/ML can provide tenants with additional confidentiality assurances, whereby the data and compute can be “sealed” from other parties and potentially the provider itself. This becomes particularly relevant where the data, computation, or results may be personal or otherwise sensitive. Moreover, integrity guarantees (over the code, data, and execution) can also be used to provide evidence demonstrating that the data was used only for particular purposes.

Data sharing and transfer. Many data processing scenarios involve the transfer of data, whether as part of its collection, aggregation, distribution, computation, or through interactions with various online services. Data may flow across technical boundaries (software, service, or hardware) or administrative domains (that is, infrastructures of different organizations).²⁴ In a cloud context, this can include data moving in-cloud (for example, across VMs, containers, and other hosted services), as well as data moving to/from cloud services through the tenants up/downloading data, and interactions with remote services.

The remote-attestation functionality of enclaves can be relevant here to help

ensure the confidentiality and integrity of code, data, and computation. This is because remote attestation allows an interrogation of the machine to which data is being considered for transfer by verifying, for example, that it is correctly configured, runs the correct code or applications, is operated by the expected organization, and so on. This can provide some assurance that data will be processed in an appropriate and expected manner. In this way, the technology can allow for more measured and controlled data transfers between parties.

From a data protection perspective, remote attestation can give controllers a greater degree of control over the processing operations for which they are ultimately responsible. Processors (typically providers) can process data only on behalf of and on the instruction of controllers (typically tenants) (GDPR Art.28 and 29), while also employing measures to ensure they process data in a manner that complies with GDPR (GDPR Art.28). Remote attestation potentially assists with this, paving a way for controllers to establish processing boundaries and to verify processor adherence. Attestation may also aid processors in demonstrating their own compliance and assisting their audit (GDPR Art.28), for example, where a provider involves others as part of its supply chain.

As such, TEEs represent one possible measure (of many) that could help controllers ensure that processors are operating appropriately. They also assist processors in demonstrating that they are meeting their obligations.

Benefits and risk. Cloud services are used by tenants for a wide variety of purposes. Given the growing awareness of legal obligations relating to data and its processing, various materials describe TEE technology as a means for tenants to meet their GDPR obligations (as mentioned earlier). Yet, while industry focuses on the benefits for tenants, less discussed is how the technology may actually serve the providers' interests.

Cloud provider liabilities. Under GDPR, the controller is the entity that remains ultimately responsible for compliance (GDPR Art.5(2)). In the cloud this is generally (but not always) the tenant. GDPR provides several ways, however, in which processors

(usually cloud providers) may be liable for certain violations, including in relation to security obligations, and therefore may face penalties for noncompliance. It follows that providers can benefit through means helping them manage these concerns.

Again, GDPR requires processors to take appropriate technical and organizational security measures, and to assist the controller in ensuring compliance with those obligations (GDPR Art.28(3)(c), Art.28(3)(f)). GDPR also requires that controllers and processors establish contractual provisions that oblige processors to make available to controllers the information for demonstrating compliance and facilitating audits (GDPR Art.28(3)(h)). Indeed, prominent cloud providers incorporate such terms into their standard form services agreements (see, for example, those for Amazon AWS² and Microsoft Azure¹⁴), although these raise practical considerations (for example, the limitations on the ability of tenants to meaningfully “inspect” a provider's operations¹⁷). This means that while both tenants and providers may be accountable to regulators and other oversight bodies for failures to meet these security requirements (which can result in potentially severe penalties), cloud providers should, according to GDPR, also be accountable in contract to the tenant and may face action, potentially including for damages, when they do not meet their obligations.

In other words, cloud providers have a number of obligations, and this exposes them to legal and financial liability. Cloud providers tend to be the major player in a technology supply chain, and as large organizations with substantial resources at their disposal with which to cover damages, they are often a prime target for litigation. Therefore, there are strong incentives for providers to limit and manage their risk exposure.

Benefits for providers. In adopting TEEs, cloud service providers may help mitigate their own liability exposure.

First, the security properties offered by enclaves can provide additional guarantees regarding the confidentiality of data, code, and processing. If there is a (plaintext) data leak, the provider may point to the use of enclaves to show that the data was protected on its platform, indicating that the responsibility for the

leak lies elsewhere. The integrity guarantees of enclaves can also help demonstrate that the cloud provider has acted appropriately (for example, in terms of system configuration, that the data was not tampered with, the right code was executed, and so on). It follows that in practice, enclaves may assist a provider by presenting evidence that it wasn't at fault. This can help in absolving liability when issues arise.

Although the provider, as a processor, still has obligations under GDPR, even where it does not know that data is personal and even if it cannot access the plaintext, the fact that the tenant requested an enclaved-backed service gives the provider an extra signal that the particular scenario may require special care. This can help inform the provider's associated risk-mitigation strategy.

Benefits for tenants. There are some benefits for tenants. First, the technology can help tenants manage their own legal obligations. As discussed, controllers can use only those processors that can provide sufficient guarantees they implement the appropriate technical and organizational measures to meet GDPR's requirements, including the security obligations (GDPR Art.28(1)). Tenants using the services of a cloud provider that employs TEEs can go some distance toward meeting their own obligations under GDPR and mitigate their own liability exposure. This is where the technology gives extra assurances that better justify the use of the cloud, while allowing tenants to benefit from the provider's functionality and expertise, where the provider, indeed, may well be better placed to deal with security and data-management concerns.

Enclaves can also provide a foundation for tenants to "audit" the provider. For example, if a cloud provider is claiming to use TEEs, the tenant should be able to verify this, as well as the associated operations and functionality that the enclave underpins. Where this is not occurring, the tenant can potentially take action against the provider.

More generally, however, enclaves, as a means for limiting a provider's involvement in processing operations (that is, by segregating data and compute), may actually encourage the uptake of the cloud in situations that otherwise would have been too risky

(for example, where data is particularly sensitive). Therefore, the technology can make cloud services more viable for prospective tenants by influencing the risk calculus such that compliance becomes less of a barrier for cloud adoption. This may be desirable for prospective customers, which may seek to migrate to the cloud to benefit from reduced operational costs (compared with running operations in-house) and to leverage the provider's other security measures, resources, and expertise.

A gain for providers. In short, TEEs can be useful for mitigating the provider's own risk. At the same time, the technology can also encourage more cloud adoption, which again benefits the cloud provider. This is interesting, given that providers are actively marketing the technology to prospective tenants, often at a premium, despite the significant benefits for providers in terms of their own risk management and in driving further business.

The power of dynamics of adoption.

A vast array of applications, services, and organizations rely on the use of cloud services. At the same time there is much consolidation among vendors in the cloud computing space, and in the tech industry more generally, in which there are a few dominant players.⁶

TEEs have the potential to help entrench the reliance on the cloud and strengthen the position of the major cloud service providers. This is because the guarantees provided by enclaves could drive further cloud adoption, including in situations where this would previously have been untenable. Moreover, if TEEs become widely accepted as best practice, it will be these providers that are best placed to meet the requisite standards. This is because of their significant resources, access to security expertise, experience with the technology, and ability to absorb the cost of implementing new technologies.

In this way, regulatory and market incentives may interact with the development and adoption of TEEs to help drive the consolidation of the technical infrastructure, which supports a range of applications, around a few dominant companies. This raises a number of concerns.⁶

First are those around security and resilience. An issue or vulnerability in a

cloud provider's service, such as an incorrectly configured TEE, can potentially impact their entire customer base. This is particularly problematic in a consolidated environment, given that the cloud aims at scale—a provider serves a vast range of customers, covering a vast range of applications (of varying importance for individuals, business, and society), meaning any issue can have systemic implications.

Further, by controlling technical infrastructure, providers have the power to act as gatekeepers, determining how, when, and why the cloud services that now underpin many applications are used. Since the major cloud providers offer a substantial collection of other products and services, some of those applications and services offered by tenants will be in competition with the offerings from the companies also providing the cloud. The power to determine who can use their services, and for what purposes, means they might—intentionally or otherwise—influence the behavior of their competitors in other markets, while at the same time generally influencing the broader application landscape.

The context for this discussion is that the dominant tech firms are in position to derive significant market and societal power.³⁰ As societal reliance on cloud services increases, so too does the power conferred on providers. As these providers were already dominant in their markets, this would, in effect, be a circular, self-reinforcing phenomenon: TEEs provide incentives for greater use of the cloud; TEEs become more widely accepted as best practice; cloud providers are best placed to implement TEEs, given their expertise and economies of scale; which provides further incentives for the use of cloud services; and so on. As a result, the current structural conditions of the cloud ecosystem, and the power of providers within that ecosystem and in society more generally, would likely be reinforced and entrenched.

Therefore, the potential for TEEs to increase the power of cloud providers and provide incentives to use the cloud should be factored into discussions concerning the dominance of technology firms—particularly in light of society's increasing reliance on technical infrastructure.³⁰

Enclaves: A basis for compliance and accountability? A related area warranting consideration is how enclaves might provide a technical foundation for enabling new measures that aim explicitly at increasing levels of assurance, compliance, and accountability²⁴—that is, to explore whether and how TEEs can provide a suitable processing environment and other guarantees upon which specific legal compliance and accountability mechanisms could be built. For example, TEEs could provide the basis for raising levels of assurance in systems audit data (evidence) through improved recordkeeping regimes (for example, by improving the security and integrity of logs and logging mechanisms¹⁰). Generally, there is a need for further work on enabling more reliable and robust mechanisms supporting the oversight, monitoring, management, scrutiny, and review of technical infrastructure and organizational practices. This is an area ripe for attention.

Concluding Remarks

Organizations that deal in data are increasingly subject to legal and regulatory obligations. In exploring how data protection relates to enclaves in a cloud context, we have shown that—despite material suggesting the possibility—TEE's will not *solve* compliance issues. They do not provide a means by which those designing, deploying, or using technical systems can avoid their data protection responsibilities. Instead, TEEs represent yet another security tool (but one tool among many) that can aid risk management. They have the potential to help in building more secure systems, and in doing so can perhaps assist with meeting *some* legal obligations.

For tenants, the real gains appear less about the legal aspect and more about reducing the compliance barriers for cloud adoption—where cloud uptake can result in cost savings. While enclaves are marketed to prospective tenants, many of the potential benefits of TEEs actually fall to the providers. That is, providers can use the technology to manage their own risks better and to drive further business. Indeed, this accords with the prior criticisms of such technology as primarily being in the interests of vendors. As outlined here, lower-level security technologies such as TEEs can

reinforce the power of the few dominant tech companies and, therefore, warrant further consideration.

There also appears to be an opportunity for the technology to support stronger governance regimes by underpinning compliance, reporting, and other accountability mechanisms. This is an area that is so far underexplored.

Overall, the above tech-legal analysis shows that there is more to enclaves than just their functionality. As such, the technology should be on the radar of not only technologists, but also businesses, regulators, policymakers, and civil society.

Acknowledgments

The Compliant and Accountable Systems Group acknowledges the financial support of the Engineering and Physical Sciences Research Council, University of Cambridge, through the Trust & Technology Initiative, and Microsoft through the Microsoft Cloud Computing Research Centre. Do Le Quoc is supported by funding from an innovation program under the LE-GaTO Project (legato-project.eu), grant agreement no. 780681. C

References

1. Alder, F., Asokan, N., Kurnikov, A., Paverd, A., Steiner, M. S-FaaS: Trustworthy and accountable function-as-a-service using Intel SGX. In *Proceedings of the ACM SIGSAC Conf. Cloud Computing Security*, 2018, 185–199; <https://dl.acm.org/doi/10.1145/3338466.3358916>.
2. Amazon. AWS GDPR Data Processing Addendum; https://d1.awsstatic.com/legal/aws-gdpr/AWS_GDPR_DPA.pdf.
3. Anati, I., Gueron, S., Johnson, S., Scarlata, V. Innovative technology for CPU-based attestation and sealing. In *Proceedings of the 2nd Intern. Workshop on Hardware and Architectural Support for Security and Privacy*. ACM.
4. Anderson, R. Cryptography and competition policy: issues with “trusted computing.” In *Proceedings of the 22nd Annual Symp. Principles of Distributed Computing*, 2003, 3–10.
5. Arnaudov, S. et al. SCONe: Secure Linux Containers with Intel SGX. In *Proceedings of the 12th Usenix Symp. Operating Systems Design and Implementation*, 2016, 689–703; <https://dl.acm.org/doi/10.5555/3026877.3026930>.
6. Cobbe, J., Norval, C., Singh, J. What lies beneath: Transparency in online service supply chains. *J. Cyber Policy* 5, 1 (2020), 65–93.
7. European Commission. What is personal data? https://ec.europa.eu/info/law/law-topic/data-protection/reform/what-personal-data_en.
8. Gollamudi, A., Chong, S. Automatic enforcement of expressive security policies using enclaves. In *Proceedings of the ACM SIGPLAN Intern. Conf. Object-oriented Programming, Systems, Languages and Applications*, 2016, 494–513; <https://dl.acm.org/doi/10.1145/2983990.2984002>.
9. Hunt, G., Letey, G., Nightingale, E. The seven properties of highly secure devices. Microsoft Technical Report MSR-TR-2017-16; <https://www.microsoft.com/en-us/research/publication/seven-properties-1st-edition/>.
10. Karande, V., Bauman, E., Lin, Z., Khan, L. SGX-Log: Securing system logs with SGX. In *Proceedings of the 2017 ACM Asia Conf. Computer and Communications Security*, 19–30.

11. Le Quoc, D., Gregor, F., Arnaudov, S., Kunkeland, R., Bhatotia, P., Fetzter, C. secureTF: A secure TensorFlow framework. In *Proceedings of the 21st Intern. ACM/IFIP Middleware Conference*, 2020.
12. Le Quoc, D., Gregor, F., Singh, J., Fetzter, C. SGX-PySpark: Secure distributed data analytics. In *Proceedings of WWW '19: The World Wide Web Conf.*; <https://dl.acm.org/doi/10.1145/3308558.3314129>.
13. Linux Foundation. Confidential Computing Consortium, 2020; <https://confidentialcomputing.io/>.
14. Microsoft. Online Services Data Protection Addendum; <https://www.microsoft.com/en-us/licensing/product-licensing/products>.
15. Microsoft. Virtualization-based security (VBS) memory enclaves: Data protection through isolation, 2018; <http://bit.ly/3ps8rF2>.
16. Microsoft. Azure confidential computing, 2020; <https://azure.microsoft.com/en-us/solutions/confidential-compute>.
17. Millard, C.J., ed. *Cloud Computing Law*, second edition. Oxford University Press, 2021.
18. Nilsson, A., Bideh, P.N., Brorsson, J. 2020. A survey of published attacks on Intel SGX; arXiv:2006.13598.
19. Ohrimenko, O., Schuster, F., Fournet, C., Mehta, A., Nowozin, S., Vaswani, K., Costa, M. Oblivious multi-party machine learning on trusted processors. In *Proceedings of the 26th Usenix Conf. Security Symposium*, 2016; <https://dl.acm.org/doi/10.5555/3241094.3241143>.
20. Open Enclave SDK. 2019; <https://github.com/openenclave/openenclave>.
21. Schoen, S.D. Trusted computing: promise and risk. *Electronic Frontier Foundation*, 2003; https://www.eff.org/files/20031001_tc.pdf.
22. Schuster, F., Costa, M., Fournet, C., Gkantsidis, C., Peinado, M., Mainar-Ruiz, G., Russinovich, M. VC3: Trustworthy data analytics in the cloud using SGX. In *Proceedings of the 2015 IEEE Symp. Security and Privacy*; <https://ieeexplore.ieee.org/document/7163017>.
23. Schwarz, M., Weiser, S., Gruss, D. Practical enclave malware with Intel SGX. In *Detection of Intrusions and Malware, and Vulnerability Assessment*. Springer International Publishing, 2019, 177–196.
24. Singh, J., Millard, C., Reed, J., Cobbe, J., Crowcroft, J. Accountability in the IoT: Systems, law, and ways forward. *IEEE Computer* 51, 7 (2018), 54–65; <https://ieeexplore.ieee.org/document/8423131>.
25. Skillern, R. Intel architecture enables new IBM cloud service with enhanced container security. Intel IT Peer Network, 2018; <https://itpeernetwork.intel.com/intel-ibm-cloud-container-security>.
26. Tarkhani, Z., Madhavapeddy, A. Sirius: enabling system-wide isolation for trusted execution environments, 2020; arXiv:2009.01869.
27. Tsai, C.-C., Porter, D. E., Vij, M. Graphene-SGX: A practical library OS for unmodified applications on SGX. In *Proceedings of the 2017 Usenix Annual Technical Conf.*, 645–658; <https://dl.acm.org/doi/10.5555/3154690.3154752>.
28. U.S. Department of Defense. 1985. DoD Trusted Computer System Evaluation Criteria.
29. Van Bulck, J., Oswald, D., Marin, E., Aldoseri, A., Garcia, F.D., Piessens, F. A tale of two worlds: Assessing the vulnerability of enclave shielding runtimes. In *Proceedings of the 2019 ACM SIGSAC Conf. Computer and Communications Security*, 1741–1758; <https://dl.acm.org/doi/10.1145/3319535.3363206>.
30. van Dijk, J., Poell, T., de Waal, M. *The Platform Society: Public Values in a Connective World*. Oxford University Press, 2018.

Jatinder Singh heads the Compliant and Accountable Systems Research Group in the Department of Computer Science and Technology, the University of Cambridge, U.K.

Jennifer Cobbe is a research associate in the Compliant and Accountable Systems Group, Department of Computer Science and Technology, the University of Cambridge, U.K.

Do Le Quoc is a postdoc at the Systems Engineering Group of TU Dresden, Germany, and a co-founder of Scontain UG.

Zahra Tarkhani is a Ph.D. candidate in the Systems Research Group at the Cambridge University Computer Laboratory, Cambridge, U.K.

Copyright held by authors/owners.