intertrust®

White Paper

Platform

Maintaining data transparency for compliance with data privacy regulations

Contents

On May 25th, 2018, the GDPR (General Data Protection Regulation) came into effect. While the GDPR is an EU (European Union) regulation, the GDPR applies to the many international companies operating in the EU and has influenced a number of data privacy regulations around the world, including the California Consumer Privacy Act (CCPA). With fines that can range up to 4% of a company's annual revenue or €20 million (approximately $24 million), the GDPR has singularly focused corporate attention on ensuring that their data operations comply with applicable data privacy regulations. One of the core themes behind the GDPR and similar data privacy regimes is the notion of data transparency. This paper will discuss the main principles of data transparency and the costs that come when these aren't followed. We will also look at the limitations of some technical compliance approaches and a solution that can overcome these.

What is data transparency?

People may be familiar with the concept of privacy notices, the oft-mentioned but little read data privacy legal notices that apps and websites make available to their users.

However, ask someone if they actually know what a company is doing with their data and very few, if any, will be able to answer. This is the exact opposite of data transparency. We are defining data transparency as the capability of an organization, in response to a legitimate enquiry from a user or regulatory agency, to clearly demonstrate how data is being used.

To explain this further, let's broadly look at some of the basic principles of the GDPR that inform the regulation's approach to data transparency.

Data subjects and their rights

Data subjects and their rights form the core of the GDPR. To start with, a data subject is pretty easy to define, it's an actual individual human being. A further modifier is that the individual can be identified by data that an organization holds. The GDPR gives individuals eight rights:

1. Know that an organization is handling (usually referred to as processing) their data and providing access to it upon request.
2. Correct personal data if it's inaccurate.
3. Ask that certain data be deleted even if it has been made public.
4. Restrict how data is being handled.
5. Receive clear information that their data is being handled.
6. Move data between organizations if so desired.
7. Object to their data being handled.
8. Any decisions made using their data can't be solely done by automation.

Consent management

The GDPR gives data subjects broad and sweeping rights. Allowing a data subject to know that an organization is handling their data is interpreted to mean that the organization has to receive express permission from a data subject to do so. Relying on a passive method such as displaying a link to a privacy notice will not work. For organizations, a concrete result of this right is they must track these permissions. To complicate things even further, this permission must extend to - and tracked across - contractors and/or partners who are given access to the data subject's data.

Privacy By Design

Quite often, software systems designed to handle personal data are criticized as adding on data privacy measures later on in the development process. This can make them difficult and/or expensive to implement. Privacy by Design as a concept predates the GDPR and at its core means that organizations should include privacy considerations into the creation of technology and processes that handle personal data from the beginning of the design and development process. This should also involve all the sections of the organization that interact with personal data.

Data localization

This is a regulatory requirement that states data must be physically retained within a certain geographic area. While the GDPR doesn't require organizations holding data from EU citizens to keep that data within the EU, should that organization wish to transfer that data to another country, they can only do so if the receiving country's data privacy regulations are certified to be equivalent to the GDPR. There are also other localities around the world that have specific data localization requirements.

How could lack of data transparency affect the bottom line?

From a practical perspective, for organizations to comply with the GDPR and similar data privacy regulations, data transparency means being able to concretely demonstrate that their software systems that handle sensitive data and the organizations that interact with the same follow these principles.

To do this, these same systems must be able to track and report such facts as where the data exists and who has access to it. It is clear that ensuring data transparency is necessary to comply with the GDPR and other data privacy regulations. Beyond avoiding fines, can implementing a data transparency regime help with the bottom line?

Let's look at a very real threat to any organization's bottom line - data breaches. In today's connected world, unfortunately the threat of a data breach hangs over most any organization.

According to the security software firm Varonis, in 2020 there were a total of 3950 verified data breaches in 2020.[1] In a joint report from IBM and the Ponemon Institute, between May 2020 and March 2021, the average total cost associated with a data breach was $4.62 million.[2] The same report also points out that personally identifiable information (PII), the type of information covered by data privacy regulations, were the most commonly stolen type of data with 44% of breaches including the loss of this type of record. On top of that, the cost associated with a single PII record, $180, was the highest of any record type.

Figure 2
Beyond the reputational damage, data breaches can have a major impact on an organization's bottom line. Data is from Varonis[1] and IBM/Ponemon Research[2].

3950	44%	$180	$4.62M	$5.65M
verified data breaches in 2020[1]	of stolen records was personally identifiable information[2]	Cost associated with one stolen personally identifiable information record[2]	Average total cost of a data breach between May 2020 and March 2021[2]	Average total cost of a data breach between May 2020 and March 2021 in a highly regulated industry sector[2]

Limitations of common approaches

Of course, organizations have already developed systems to comply with the data transparency requirements of data privacy regulations. With some of the common ways that these systems have been implemented, there are some limitations.

Siloed data architectures

One approach is to take regulated PII and place it in a silo where access is tightly controlled. This may work for an application environment for which the data silo was originally developed, but as applications change or new applications are brought on, the original data silo may no longer fit the needs of the changing organization. The organization may have to duplicate the same data in other silos for these new applications, thus increasing the chances for it to be improperly accessed.

Limited application access transparency

Traditional identity service providers[3] are focused on controlling human access to applications and data, and generally do a good job of providing an audit trail for regulatory compliance. That being said, if a regulator asks an organization to show an audit trail for the third party software applications that have access to PII, these providers will not be much help. The organization must then work with whatever audit feature is made available by each of the applications that the software application providers include.

Siloing data by geography

In response to regulations requiring that personal data remains within a certain geographic area, organizations are forced to silo off some of their data to make sure it is physically stored in the area. For organizations that have applications that span multiple geographic areas, this data siloing complicates the integration of data.

Uncontrolled profiling

Some traditional identity service providers may not have data access controls that cover both individuals and third party software applications. This can open up access to data to applications which could potentially be passing on data and profiles to unauthorized parties.

Incomplete masking and anonymization techniques

The GDPR and other data privacy regulations require that organizations use certain techniques to protect personal data. One of these is data masking where certain parts of data, such as email addresses or telephone numbers, need to be obfuscated in many situations. Another one is data anonymization or pseudonymization where technical measures are used to make data analysis useful while reducing the ability to mine it for personally identifiable information. Organizations often have adopted data operations tools which may not have adequate solutions for these techniques.

Data rights management: one potential solution

Data rights management platforms currently on the market can help ensure data transparency as well as reduce data breaches and their cost.

The IBM/Ponemon Institute report referenced above notes that organizations operating in highly regulated business sectors such as energy, health care, finance, education, etc. had a higher average cost associated with breaches in the $5.65 million average range. The report attributes the difference to an increased level of fines, penalties and lawsuits. Since a data transparency system backed by a data rights management platform will help demonstrate regulatory compliance, it could help either avoid or reduce the fines and penalties associated with a data breach.

Another factor is that effective data privacy protection requires that both human and machine access to PII records be limited to people and systems that can demonstrate a need to access the data. A data transparency investigation should be able to demonstrate that this principle was followed. Updated identity service providers that are part of data rights management platforms should be used to authenticate both humans and software data access following data privacy policies. This can reduce costs associated with data breaches by limiting the number of people and systems that can access PII, reducing avenues for unauthorized data access. Should a data breach occur, this service can also help in the forensic analysis of how the data breach occurred.

The software access control part of an identity service is an important element in preventing data breaches. Organizations focus on controlling human access to PII but often also depend on third-party analytics programs to process the same data. These processes bring the potential of either advertently or inadvertently exposing this sensitive data. An updated identity service provider can isolate these programs within the organization's system and control both access to data sources as well as where the analytics results are sent to. Accordingly, personal data can be both tightly controlled and doesn't need to leave the system to accommodate third party programs.

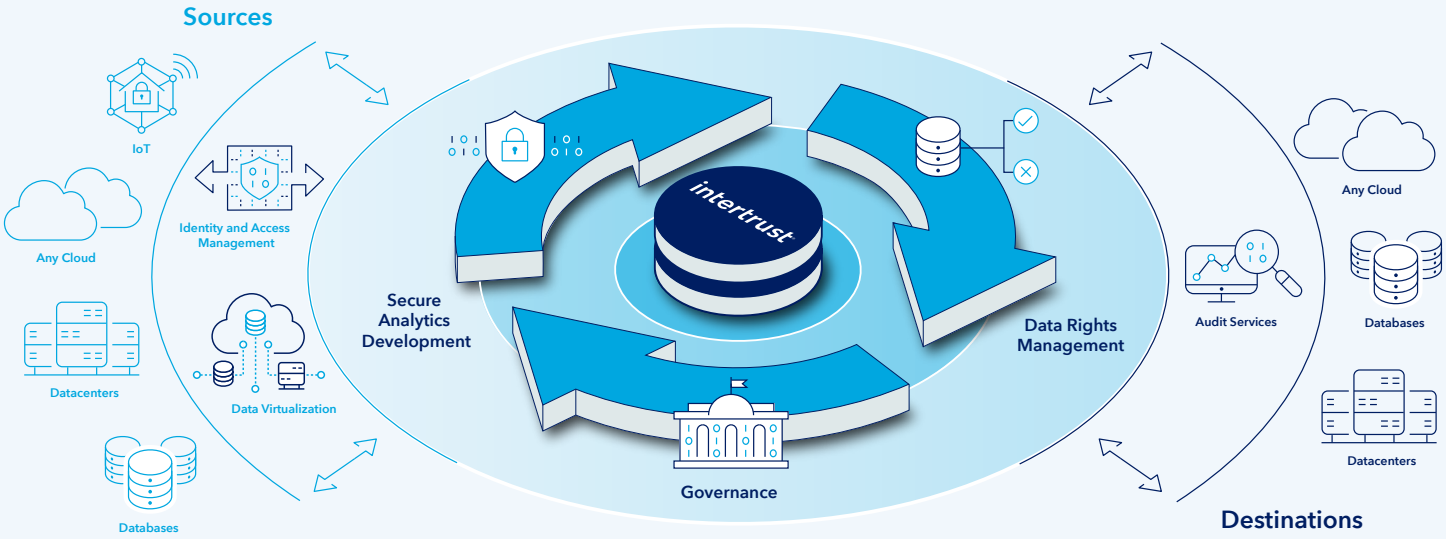
Another way that data rights management systems can help reduce the "attack surface" that allow data breaches to occur is by avoiding unnecessary duplication of personal data. Often data operations systems are set up in such a manner that to perform analytics, the system design requires that data be moved from its original location to another location such as a data lake or data warehouse. Each time data is duplicated, it increases the potential for bad actors to access that data, such as by increasing the chances of misconfigured cloud implementations.

The IBM/Ponemon report states that cloud misconfiguration is the 3rd largest attack vector resulting in data breaches coming in at 15% of breaches. The average cost of a data breach caused by a cloud misconfiguration was $3.86 million. A data rights management system that includes a data virtualization feature can avoid this issue by allowing a system to query data from its current location for analytics use. There is no need to duplicate the data for moving into either in-house or external systems.

Introducing Intertrust Platform

There are a number of data privacy centric solutions that claim to be able to help organizations meet their data privacy regulatory requirements. Still, these will need to be integrated into an organization's data operations system. Another option to consider is selecting a data operations platform based on data rights management capabilities. One such solution is Intertrust Platform™.

Intertrust Platform is a versatile data operations platform that addresses organizations' needs for secure collaboration and data interoperability. It is also designed to be a robust solution for data transparency. Backed by Intertrust, a company with over 30 years of history in protecting valuable distributed data for major international companies, the Platform is being used by major international companies to manage distributed data and devices at scale, so they can make data-driven business decisions, confidently and securely while complying with data privacy regulations.



The Platform supports three major functions:

Identity and access management

The Platform enables data governance via fine-grained user authentication via open standards. This feature can allow organizations to fine-tune access privileges to sensitive data for both internal and external users and is an important feature for data transparency and regulatory compliance.

Data virtualization

Data does not need to be migrated onto Intertrust Platform. Fine-grained privileges may be applied to existing datasets, regardless of their location, creating a unified point of access control that governs all interactions via the Platform's data interfaces. The Platform also enables users to create 'virtual datasets' by joining data from one or more physical data stores.

Secure execution

A key feature of the updated identity service provider Intertrust Platform provides workflow environments where containerized workloads can be deployed and executed in a managed cluster across any cloud/on-premises setup. Both ingress and egress to containers are constrained by network policies. When coupled with the Platform's governance mechanisms, this ensures that workloads may only access data to which privileges have been granted and that data may not transit beyond the governance boundaries.

For auditing purposes, the Platform also provides a robust logging mechanism that meets stringent standards, generating a secure log that is controlled, immutable, and protected.

Additionally, Intertrust Platform includes state of the art data obfuscation techniques that support compliance with GDPR and other data privacy regulations.

Data transparency and Intertrust Platform

As a complete data operations platform based on extensive data rights management technology, Intertrust Platform not only forms the underpinning of a modern corporate data operations ecosystem, it's feature can be used to improve data transparency and reduce the risk of data breaches.

The combination of the Platform's Identity and Access Management and Secure Execution Environments acts as an updated identity service provider and forms the basis for a fine-grained authentication and policy-based governance system for both human users and third-party software applications. As such, it simplifies regulatory compliance and data transparency in a number of ways. One is that by allowing for fine-grained governance, Intertrust Platform can ensure that access to personal data is restricted to the users and software applications who need it and these entities can only access the data they have permissions for. This avoids the issue of "over permissioning" where an user or software program may be given more permission than is needed to perform their function. By covering both human users and software, organizations need only work with one tool for data transparency purposes. Furthermore, combined with the auditing capabilities of the Platform, organizations can track access logs in response to regulatory requests or for breach forensics.

Data virtualization brings a number of benefits. By using a virtual layer to tie together datasets in their original sites, organizations can avoid the increased data breach potential caused by moving data or duplication of data. It also reduces the chance of unauthorized copies of datasets that can be missed in responses to regulatory requests as well as create more potential for data breaches. The data virtualization feature can help reduce data breaches in another way, and that is by simplifying the ability for organizations to use hybrid clouds. The IBM/Ponemon research indicated that the average cost of a data breach for an organization that used a hybrid cloud architecture was $3.61 million, which was less than either a purely public cloud or private cloud approach.

The Platform's data virtualization capabilities can help with data localization by ensuring that data subject to these regulations can be left in the geography and institute policies to ensure that any access is in accordance with those regulations. Intertrust also has a partnership with InCountry that helps to simplify the work of complying with data residency requirements.[4]

Conclusion

Organizations around the world are working hard to comply with data privacy regulations and their data transparency requirements.

Beyond regulatory compliance, instituting data transparency capabilities into a data operations system can help reduce the costs associated with data breaches, bringing a direct benefit to the organization's bottom line. Organizations have a number of options in choosing systems to help them meet data regulatory requirements, many of them focused specifically on data privacy. Intertrust Platform is a full-fledged data operations platform whose features can also support a robust data privacy and transparency system.

Sources

1 https://www.varonis.com/blog/data-breach-statistics/

2 *Cost of a Data Breach Report 2021*, IBM Security, Ponemon Institute

3 Identity service providers are services that maintain the identity and authentication of information and govern the process of authenticating applications and users.

4 For more information about InCountry please see https://incountry.com/

intertrust®

Building trust for the connected world.

Learn more at: intertrust.com/platform
Contact us at: +1 408 616 1600 | dataplatform@intertrust.com

Intertrust Technologies Corporation
400 N McCarthy Blvd, Suite 220, Milpitas, CA 95035