Evaluating Blockchain within the Legal Frameworks of the *General Data Protection Regulation* and the *California Consumer Privacy Act of 2018*

August 3, 2018

Edward Percarpio
Regulatory Lead, Polyphonic Labs

Konrad Rauscher
Founder, Technology Lead, Polyphonic Labs

Table of Contents

EXECUTIVE SUMMARY

This Report examines the relationship between blockchain and two recently codified data privacy regulations: the European Union's General Data Protection Regulation (GDPR) and the California Consumer Privacy Act of 2018 (CCPA). In light of new obligations imposed by these regulations (e.g. individual rights to access, delete, and port their personal data), privacy and security experts have questioned whether blockchain is able to comply with these new requirements. Skeptics note that blockchain has certain intrinsic characteristics – such as the immutability of the ledger – that prohibit the implementation of compliance requirements such as rectification or deletion. Proponents argue the opposite, stating that blockchain poses unique solutions to implementing otherwise difficult compliance requirements (e.g. verifiable erasure, 'objective' record-keeping). However, the current level of discourse around these topics often regards "blockchain" as a single immutable concept without considering how public, permissioned, and private ledgers offer unique challenges and different compliance obligations. From a legal perspective, too often does polemic jump to compliance obligations without first considering whether and how blockchain networks fit within the scope of the law itself.

When discussing blockchain/distributed ledger architectures and data protection regulations, further consideration is given to the application of these frameworks to the derivatives market. This additionally necessitates a comparative regulatory analysis of how GDPR and Know-Your-Customer (KYC) and Anti-Money Laundering (AML) requirements interact with each other and with blockchain.

This Report achieves the following tasks:

1. Describes the scope, definitions, and responsibilities of actors under the GDPR, CCPA, and American KYC-AML regulations.
2. Discusses the role of actors in a blockchain network and their legal responsibilities.
3. Examines the impact of different blockchain architectures on the legal responsibilities of its actors.
4. Discusses the ability of different distributed ledger architectures (e.g. Sovrin, Hyperledger Fabric, Teechain) to comply with regulatory requirements relevant to the derivatives market (e.g. KYC-AML, FCRA, GLBA, Dodd-Frank).
5. Provides recommendations for leveraging blockchain such that:
    a) Regulatory obligations can be met and, in certain circumstances, enhanced as opposed to traditional non-blockchain architectures.
    b) Entities can apply multiple blockchain designs (public, permissioned, private) to comply simultaneously with multiple national regulatory requirements.
    c) Security can be preserved in a post-quantum environment, while simultaneously maintaining regulatory transparency and accountability.

1 Introduction

The General Data Protection Regulation (GDPR) is a European Union regulation that offers EU citizens increased rights over their personal data, such as the right to access, port, rectify, and delete this data. Since the adoption of the GDPR in 2016, countries around the world have scrambled to implement similar principles in order to ensure that data remains able to flow smoothly across borders.[1] Notable countries that have made legal movements towards GDPR-oriented legislation include Brazil (LGPD), Japan (APPI), South Korea (PIPA), and Canada (PIPEDA). Some (e.g. Brazil) have implemented new legislation entirely, while others are updating existing legislation to maintain an "adequacy" designation (e.g. Canada).

While the US has yet to reach a national determination of whether it will toe the line, California recently passed a law granting its residents data privacy rights very similar to those afforded by the GDPR. Slated to come into effect January 1, 2020, the *California Consumer Privacy Act of 2018* (hereafter referred to as CCPA), puts pressure directly on the heart of the American tech industry to build technologies that implement these rights, or otherwise suffer strict penalties from both Californian and European regulators.

This shifting regulatory environment has prompted discussion around whether and how blockchain fits into these new regulatory requirements. Many pundits feel that blockchain is inherently incompatible with certain GDPR principles because of its immutability. From the perspective of these commentators, if an individual wishes to exercise their right to rectify or delete their personal data, companies that use blockchain technology to record personal data will lack the ability to remove or alter it and thereby fail to comply with two critical GDPR principles.[2] Others worry that even if one were to encrypt personal data on-chain as a potential solution for erasure,[3] blockchain's immutability makes it impossible for entities to update their encryption standards as existing methods of encryption become insecure with advances in processing power and decryption technologies.

While these arguments raise valid points about the potential conflict of certain blockchain constructs with current privacy laws, they often fail to discuss critical nuances in legal scope and requirements, and in different blockchain designs. This failure leads to myopic or misleading ideas of where/how blockchain and the GDPR help or hurt one another.[4] This report aims to provide a nuanced and thorough examination of the GDPR/CCPA as they pertain to blockchain, and provide recommendation for how blockchain could be leveraged to support certain GDPR principles and requirements.

[1] Chapter 5 of the GDPR restricts transfers of EU residents' personal data unless the country or entity has similarly robust legal safeguards in place. For more information, see the Annex.

[2] For examples of this concern in popular discourse, see here, here, or here.

[3] This will be discussed later in the report, but the essential idea is that an entity can implement effective data "erasure" by throwing away the encryption key, which makes the encrypted data inaccessible.

[4] E.g. the (mistaken) belief that personal data must be written on-chain, or the (mistaken) belief that the GDPR restricts personal data from leaving the EU, which would cause problems with blockchain's decentralized ledger system.

Before analyzing the GDPR/CCPA and blockchain as either symbiotic or conflicting constructs, it is essential that we first outline the parameters that constrain the scope of each law and bound the space in which we will discuss blockchain. This includes defining 1) the actors and their responsibilities in each law and in blockchain networks, 2) the types of personal data that each law regulates, and 3) the geographic/operational boundaries to which the laws restrict themselves. Then we can apply "blockchain" – or different blockchain manifestations and applications – to these legal structures and requirements to see how they impact or conflict with each other.

A final note regarding the legal scope of the Report: since the object of this Report is to contrast technological affordances/limitations with legal requirements, we will restrict our discussion of the GDPR/CCPA to the particular aspects of the laws that impose these requirements. This means that we will focus mostly on pertinent data subject rights (GDPR Arts. 15-20, CCPA Sec. 1798.100, 1798.105), business obligations (GDPR Arts 24-30), security and data protection by design requirements (GDPR Arts. 25 & 32), and certification standards (GDPR Arts. 42 & 43). This is not to say that other requirements (e.g. transparency obligations, data breach protocol) are less important; they just happen to be less relevant to a discussion involving blockchain. However, we do provide an Annex at the end of the Report that describes other critical GDPR requirements that fell beyond the scope of the Report.

2 Roles & Responsibilities

This section outlines the actors in each legal system, contrasting the roles, responsibilities, and liability associated with legal roles and the roles we find in common blockchain architectures.

2.1 GDPR

There are three main types of actors in the GDPR, not including the regulators (called Data Protection Authorities, or DPAs). Each actor is distinguished by the way they relate to the personal data being processed.

2.1.1 Data Subjects

a. Role

Data subjects are the "identified or identifiable" people who the personal data is about (Art. 4(1)).

b. Responsibilities

They are responsible for understanding the ramifications of disclosing their data or using a service that requires personal data disclosure. They are also responsible for exercising their rights as a data subject under the GDPR, which includes the right to:

- **Access** the personal data an organization holds about them (Art. 15).
- **Rectify** any inaccurate data about them (Art. 16).
- **Erase** their personal data if it is no longer necessary in relation to the purpose for which it was collected, or if the business no longer has a legal ground for processing (e.g. if the customer withdraws consent) (Art. 17).
- **Restrict** the processing of personal data in the event that its accuracy or legal grounds for processing are contested (Art. 18).
- **Port** their personal data; i.e. to receive it in a structured and machine-readable format and transmit it to another organization (Art. 20).[5]
- **Object** to the processing of personal data (Art. 21).[6]

Finally, data subjects are the initiators of the accountability process. The GDPR accountability mechanism is structured in such a way that unless a DPA decides to audit a business (likely because it is a large company known to process terabytes of personal data), the only way the DPA will be informed of potential noncompliance is if a data subject submits a complaint to the DPA, or if the company reports a data breach and the DPA decides to audit.

[5] The right to portability applies only under instances where processing is both based on consent and carried out by automated means (Art. 20(1)(a)(b)).

[6] The right to object only applies in instances where the processing is used for direct marketing (Art. 21(2)) or if processing is based on the legal grounds of legitimate interests or public interest (Art. 21(1)).

2.1.2 Data Controllers

a. Role

A data controller determines the purposes and manner in which personal data should be processed (Art. 4(7)). This includes:

- **What** personal data is processed
- **How** personal data should be processed
- **Who** will process the personal data
- **Where** personal data will be stored
- **To whom** personal data will be disclosed

b. Responsibilities

Since controllers make these decisions, they are responsible for ensuring that data collection, processing, retention, and disclosure abide by GDPR requirements. Their obligations include:

- Ensuring that personal data has been obtained correctly (e.g. gaining explicit consent rather than bundling consent in large privacy policies and terms of service documents).
- Providing the appropriate legal justifications for each processing activity.[7]
- Keeping records of data processing activities, as well as how each activity is secured and justified (Art. 30).
- Ensuring that all subcontractors – "data processors" – are compliant (Art. 28(1)) and bound to a service agreement that includes contract clauses required under the GDPR (Art. 28(2)(3)).
- Notifying both the data subject and the DPA in the event of a data breach (Arts. 33 & 34).
- Performing Data Protection Impact Assessments (DPIAs) for any new technologies that process personal data in a manner that might pose a risk to the data subject (Art 35).
- Providing transparent communication to the data subject (Arts. 12-14), explaining:
    - The types of personal data processed, the purpose for processing, the recipients of the data, and (if applicable) the existence of any extra-EEA transfers and how personal data will be protected in such instances.
    - Data subject rights under the GDPR, and how they can exercise them.
    - Any automated decision-making (e.g. profiling), and what the logic and potential consequences of the decision may be (Arts. 13(1)(f) and 14(2)(g)).

In some cases, controllers make data processing decisions in tandem. In this case, they become "**joint controllers**" and share the obligations between them. These obligations, and liability, should be determined in a contractual agreement made between controllers prior to engaging in any data processing activities.

2.1.3 Data Processors

[7] This is called "lawful basis for processing". There are six lawful bases, which can be found under GDPR Article 6.

a. Role

Data processors process data on behalf of the data controller (Art. 4(8)). The distinction between the two lies in the decision-making power. Controllers define the parameters by which the processor is allowed to process data.

b. Responsibilities

Processors are therefore responsible for:

- Recording all processing activities.[8]
- Demonstrating their GDPR compliance to data controllers, by ensuring they have the legal, technical, and organizational measures in place to sufficiently comply with GDPR requirements. These include:
    - Proving that their systems follow principles of data protection by design and default.
    - Demonstrating that appropriate legal contracts are in place if they intent to transfer personal data outside the EEA.
    - Proving that their systems can provide the data controller with necessary information and functionalities, such as:
        - Data access, rectification, erasure, conversion into machine-readability
        - Adequate security techniques and systems (elaborated under Art. 32)
        - Providing the data controller (and the DPA) with data breach report information within 72 hours.

2.1.4 Liability

Under the GDPR, liability distributed between the controller and processor depends primarily on their contract. This ambiguity exists because liability is ultimately based on the degree to which each party is responsible for the damage, which in turn depends on the nature of the data processing contract made between the controller and processor. Yet due to each party's role in contract-formation, it is slightly harder for processors to get caught up in the chain of accountability. This is because processors can only be liable if they act outside the parameters drawn up by the controller in their contract, or if they act outside their processor-specific obligations (e.g. record-keeping).[9] Controllers also have greater intrinsic liability risks simply because they have more responsibilities under the GDPR (see above) for which they will be held accountable regardless of the data processor contract (e.g. transparency requirements). But in any case, each party (or both) is liable for infringements unless they can prove that they "are not in any way responsible for the damage" (Art 82(3)). And again, the proof of indemnity or lack thereof rests in the roles and responsibilities established in the controller-processor contract.

[8] There is a caveat to the record-keeping requirement for both controllers and processors: if the company has fewer than 250 employees and does not process personal data on a regularly and large-scale basis, then the company does not need to keep a record.

[9] Article 82(2).

2.2 CCPA

The CCPA addresses two actors instead of three, also designated by their relationship to the personal data at hand. These actors are "consumers" and "businesses."

2.2.1 Consumers

a. Role

Consumers are similar to the GDPR's "data subject" in that they are the identified individuals to whom the personal information refers. While data subjects in the GDPR are not geographically restricted, Consumers under the CCPA refer only to California residents (see CCPA *Sec. 2*).[10]

b. Responsibilities

Similar to the GDPR, consumers have the primary responsibility to exercise control over their information and ensure accountability over their rights to their personal information. These rights include the right to:[11]

(1) **Know** what personal information is being collected, sold, or disclosed, as well as:
    a. the categories of sources from which the personal information is collected
    b. the categories of third parties to whom it will be disclosed
    c. the business or commercial purpose for collecting or selling the data.[12]
(2) **Object** to the sale of personal information.
(3) **Access** their personal information.
(4) **Nondiscrimination** from the business based on their privacy preferences.
(5) **Request the deletion** of personal information from the business's records as well as the records from any service provider (note the many caveats, footnoted).[13]
(6) **Opt out** from allowing the business to sell their personal information to third parties.[14]

[10] As we will elaborate in the next section, a data subject does not have to be an EU resident in order to be protected under the GDPR; businesses established in the EU who process personal data have to comply with the GDPR in those interactions regardless of the nationality or residence of the data subject (see Art 3(1)).

[11] CCPA Sec. 2; also see 1798.100(a), 1798.105(a).

[12] A big difference between the CCPA and the GDPR in this informational right is that the GDPR requires businesses to inform data subjects implicitly, whereas the CCPA specifies that businesses only have to provide this information "upon a verifiable request from the consumer" (see 1798.110(b)).

[13] 1798.105(a), (c). Businesses don't have to comply with the request for deletion if the customer's information is needed in order to perform the contract made with the consumer (d)(1), detect security incidents or protect against fraudulent or malicious activity (d)(2), debug to identify or repair errors that affect functionality (d)(3), exercise the business's rights or those of another consumer (d)(4), comply with the California Electronic Communications Privacy Act (d)(5), engage in research of public interest (d)(6), to enable solely internal uses that a customer could reasonably expect based on their relationship with the business (d)(7), to comply with a legal obligation (d)(8), or otherwise use the personal information in a lawful manner that is compatible with the context with which the information was provided (d)(9).

[14] 1798.120(a).

2.2.2 Businesses

a. Role

A business has a much narrower scope under the CCPA than both data controllers and data processors under the GDPR. A "business" must:

- Be a sole proprietorship, partnership, LLC, corporation, association, or other legal entity that is organized or operated for the profit or financial benefit of its shareholders or other owners.[15]
- Collect consumers' personal information, or (either alone or jointly) determine the purposes and means of processing of consumers' personal information.[16]
- Do business in the State of California.

A business also must satisfy one or more additional thresholds to be considered a "business" under the CCPA. It must either:

- Have gross annual revenues over $25 million,
- Buy, receive, sell, or share the personal information of 50,000 or more consumers, households, or devices on an annual basis,[17] or
- Derive 50 percent or more of its annual revenues from selling consumers' personal information.

b. Responsibilities

Businesses are obligated to comply with the consumer rights listed above according to the following specifications, requiring businesses to:

- Provide easily accessible and comprehensible forms for consumers to exercise their rights.[18]
- Provide at least two methods for consumers to submit requests for information, including, at minimum, a toll-free telephone number and a website address (if applicable).[19]
- Disclose and deliver required information free of charge and in a "readily useable format"[20] to the consumer within 45 days of the request.[21]

[15] 1798.140(c)(1). Note that this excludes public bodies, non-profits, and any organization that is not operated "for the profit or financial benefit of its shareholders or other owners." This is a large deviation from the GDPR, which does not make such a distinction regarding profit.

[16] Ibid. This is another huge difference from the GDPR, which essentially removes what the GDPR calls "data processors" from the scope of the CCPA.

[17] 1798.140(c)(1)(B). Note that likely refer only to Californian consumers/households/devices.

[18] 1798.135(a).

[19] 1798.130(a)(1).

[20] This imposes a requirement essentially identical to the GDPR's right to data portability (Art. 20).

[21] 1798.130(a)(2). This differs from the GDPR, which requires under Art. 12(3) for businesses to respond to requests within one month of the request being.

- Inform consumers of their privacy rights in an online privacy policy, and update that information at least once every 12 months.[22]
- Record lists of the categories of personal information that have been sold and/or disclosed about consumers in the past year.[23]
- Record a consumer's opt-out request for at least 12 months before requesting authorization to sell their personal information.[24]
- If a business chooses to de-identify a consumer for compliance (i.e. to carry out a request for deletion) or security purposes, it must implement organizational processes and technical safeguards that prohibit re-identification.[25]

2.3 Summary: Business Responsibilities under the GDPR vs. CCPA

Obligation	GDPR	CCPA
Information	Yes (different requirements)	Yes (different requirements)
Access	Yes (different requirements)	Yes (different requirements)
Deletion	Yes	Yes
Rectification	Yes	Not in CCPA
Portability	Yes	Yes
Objection	Yes	Yes (only to sale of PI)
Nondiscrimination	Somewhat (only re: consent)	Yes
Record-keeping	Yes (different requirements)	Yes (different requirements)
Perform DPIAs	Yes (for new, potentially risky technologies that process PD)	Not in CCPA
Hire/Appoint a DPO	Yes (if regular, large-scale processing of special category data)	No DPO role exists in CCPA
Hire/Appoint Representative if not established in territory	Yes (if large-scale processing of special category data and no establishment in EU	No, businesses don't need separate representative in California in any situation
Explain automated decisions	Yes, if processing is regular and could legally or similarly significantly affect DS	No
Response time to consumer/subject requests	Within 30 days	Within 45 days
Protection of "minors"	Parental consent for processing of PD required under 16 years old	Parental consent for sale of PI required under 16 years old

[22] 1798.130(5).

[23] 1798.130(c)(i)(ii).

[24] 1798.135(5).

[25] 1798.140(h)(1), (2).

3 Legal Scope

3.1 GDPR

The scope of the GDPR is unique for a national/international law in that it does not confine itself to a particular geography, type of entity, or sector (as one often sees in American regulation). Instead it ties itself to an action: the processing of personal data.

Personal data refers to any information relating to an "identified or identifiable natural person" (i.e. a "data subject").[26] Personal data can be a direct or indirect identifier of a natural person – it can be an identification number, location data, an online identifier (e.g. IP address, cookies), or any reference to a factor specific to the "physical, physiological, genetic, mental, economic, cultural or social identity" of the person. **Processing** of personal data refers to any operation performed on personal data, including but not limited to collection, adaptation, organization, storage and disclosure.

With these two definitions in mind, the GDPR applies either to organizations in the EU that process personal data, or organizations not in the EU that process the personal data of EU citizens. More formally, the GDPR's scope applies to:

- Organizations (businesses, NGOs, public authorities, etc.) "established"[27] in the EU who process personal data, regardless of the characteristics/demographics of the individual whose personal data is processed.
- Non-EU established organizations who process the personal data of persons "who are in the EU",[28] where such processing relates to either a) the offering of goods and services (not contingent on payment) or b) the monitoring[29] of their behavior within the EU.

3.2 CCPA

The CCPA, while often compared to the GDPR as America's state-level analog, is neither as comprehensive a regulation as the GDPR nor does it offer the same protections, obligations, and definitional ambiguities as those encoded in the GDPR. From a macroscopic level, the CCPA focuses primarily on consumer-business interactions, particularly the affordances that businesses must give to consumers. As a result, some of the requirements in these areas are equivalent or

[26] GPDR Article 4(1).

[27] The particular definition of what constitutes an "establishment" is ambiguous. Recital 22 emphasizes the "real and effective exercise of activity through stable arrangements." These arrangements do not necessarily have to be physical. *Weltimmo v NAIH* (C-230/14) confirms that "establishment" is "broadly and flexibly" defined such that even the presence of a single representative may be sufficient to qualify an entity as "established" in the EU.

[28] This language, taken from Article 3(2) of the GDPR, is also debated. Generally, it is interpreted as synonymous with EU residents, but it also applies to an idea of probability or "reasonable expectation" that the consumer is or could be located in the EU.

[29] Monitoring generally involves tracking individuals online, often to create profiles, particularly where these profiles are used to make decisions that analyze/predict a person's preferences and behaviors.

arguably more stringent than the GDPR,[30] while other areas fleshed out by the GDPR are omitted from the law entirely.[31] These differences hold significant ramifications both for the application of blockchain, and for companies under the misguided notion that GDPR compliance will satisfy CCPA compliance (and vice versa).

The first significant difference is the definition of **personal information** (which the GDPR calls personal data), defined as "information that identifies, relates to, describes, is capable of being associated with, or could reasonably be linked, directly or indirectly, with a particular consumer or household."[32] Some legal experts, such as Baker McKenzie partner Lothar Determann, view this definition as broader in scope than the GDPR's definition of personal data.[33] One reason for this interpretation might be the CCPA's inclusion of households into the definition, rather than just individuals. This could arguably lower the threshold of identification, as one merely needs to identify a group of people (albeit small) rather than a single person within that group.

The second significant difference in scope to the GDPR pertains to the material and territorial threshold that triggers CCPA's application. First of all, as described in Section 2.2.2, only "businesses" need to comply with the CCPA. This excludes public bodies, non-profits, and any organization that is not operated "for the profit or financial benefit of its shareholders or other owners".[34] Next, businesses need to comply with the CCPA if they collect California residents' personal information or determine the purposes and means of processing that personal information. Businesses have to further meet one or more of the following thresholds before they reach the scope and need to comply:

- Have gross revenues in excess of $25 million.
- Annually buys/receives for commercial purposes the personal information of 50,000 California residents, households, or devices.
- Derives 50 percent or more of its annual revenues from selling California residents' personal information.

These scope parameters differ largely from the GDPR. First of all, the above three bullets will likely exempt a large proportion of smaller (and arguably medium-sized) businesses outside California from needing to comply with the CCPA. Second, California residents lose their protections if they (and the business) are both outside California during this processing transaction.[35] Third, the CCPA was drafted to operate subserviently to other pre-existing sectoral laws that regulate personal data, such as the Federal Credit Reporting Act ((15 U.S.C. Sec. 1681 et seq.), Gramm-Leach-Bliley Act (GLBA, Public Law 106-102), and the Health Insurance Portability and Availability Act of 1996 (HIPAA).[36] This diverges from the GDPR, which as a

[30] For example, the CCPA requires businesses to make available to consumers at least two discrete methods for submitting subject access requests (SARs), including a toll-free number (see 1798.130(a)(1) and 1798.140(i)).

[31] E.g. data protection by design, controller v. processor obligations, cross-border transfers, certification and accountability.

[32] CCPA 1798.140(o). For the laundry list of examples of personal information, see 1798.140(o)(1)(A-K).

[33] Lothar Determann, "Analysis: the California Consumer Privacy Act of 2018". *IAPP* (July 2, 2018).

[34] 1798.140(c)(1).

[35] See 1798.45(a)(6). This is in sharp contrast to the GDPR, whereby businesses need to comply with EU residents' personal data regardless of the location of any party.

[36] See CCPA 1798.145 (c) (d) (e) (f), and 1798.196.

Regulation (as opposed to a Directive) supersedes all pre-existing national laws by EU member states except for specific internal derogations described within the law itself (such as minimum age for consent, threshold for needing to hire a data protection officer, etc.). Finally, the emphasis on collection and decision-making control could lead one to believe that an entire category of data processing entity – the data processor, who processes personal data but neither collects it from the consumer nor makes decisions regarding how it should be processed – need not comply with the CCPA. This marks a huge difference, particularly for the subsequent discussion of blockchain and the roles of each type of data entity.

3.3 Summary: Definitions and Scope under GDPR vs. CCPA

Definition	GDPR	CCPA
Personal Data/Information	Just individuals' personal data. Relevant PD does not pertain to EU residents only. Could be based on PD processed by an EU-based org.	Includes households as well as individuals. Restricts relevant PI to those of Cali residents.
Businesses Affected	Any org that processes PD of people in the EU, or any org based in the EU that processes PD to offer goods and services or monitor behavior.	Excludes non-profit businesses and public bodies. Much stricter scope based on discrete numbers of PI processed and/or revenue. (See diagrams below)
Data Processor?	Yes, covered under scope/compliance requirements.	Not in scope of CCPA. Only businesses that decide how PI is processed.
Relationship to other laws	Trumps existing national laws except in certain areas (e.g. DPO requirements, minimum age for consent).	Subservient to other national laws (e.g. HIPAA, FCRA).

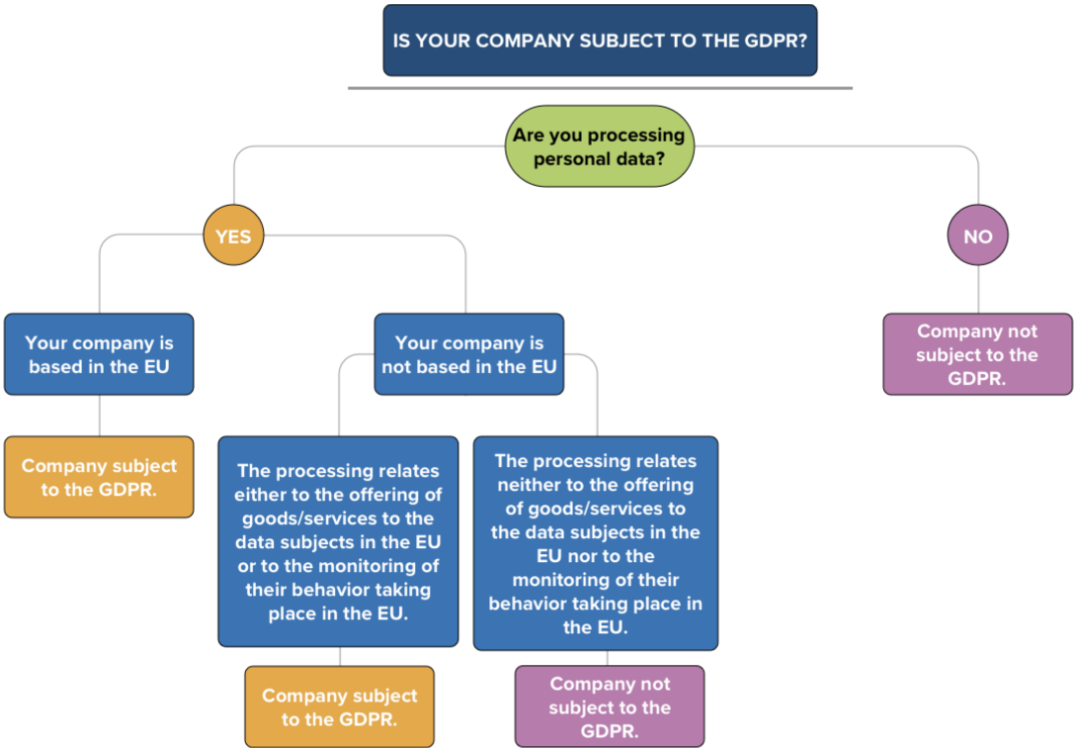

IS YOUR COMPANY SUBJECT TO THE GDPR?
Are you processing personal data?
YES
NO
Your company is based in the EU
Your company is not based in the EU
Company not subject to the GDPR.
Company subject to the GDPR.
The processing relates either to the offering of goods/services to the data subjects in the EU or to the monitoring of their behavior taking place in the EU.
The processing relates neither to the offering of goods/services to the data subjects in the EU nor to the monitoring of their behavior taking place in the EU.
Company subject to the GDPR.
Company not subject to the GDPR.

Figure 1: GDPR Compliance Scope

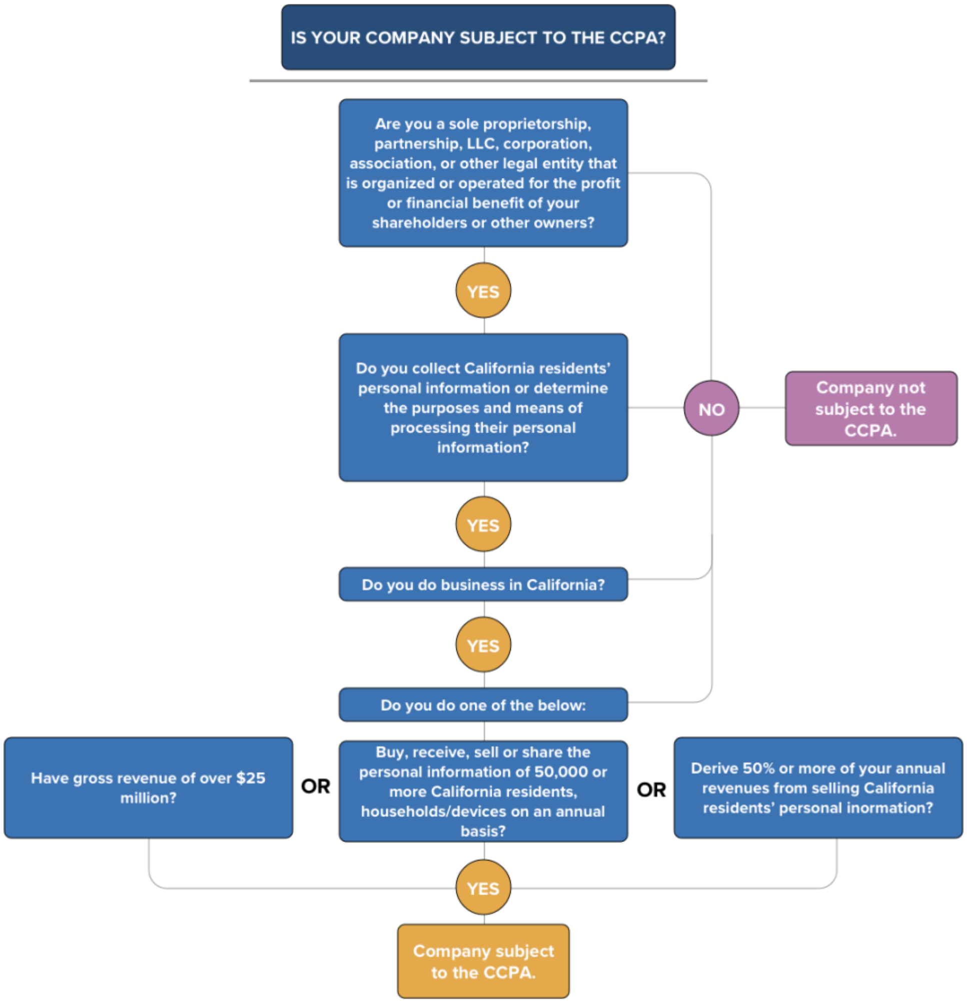

IS YOUR COMPANY SUBJECT TO THE CCPA?
Are you a sole proprietorship, partnership, LLC, corporation, association, or other legal entity that is organized or operated for the profit or financial benefit of your shareholders or other owners?
YES
Do you collect California residents' personal information or determine the purposes and means of processing their personal information?
NO
Company not subject to the CCPA.
YES
Do you do business in California?
YES
Do you do one of the below:
Have gross revenue of over $25 million?
OR
Buy, receive, sell or share the personal information of 50,000 or more California residents, households/devices on an annual basis?
OR
Derive 50% or more of your annual revenues from selling California residents' personal inormation?
YES
Company subject to the CCPA.

Figure 2: CCPA Compliance Scope

4 Further Applicable Regulation

Since blockchain networks are being applied in a number of different markets, it is important to note that compliance with the GDPR and CCPA does not exempt a business or framework from compliance with other regulations. Particularly in the U.S., federal sectoral regulations supersede the CCPA. This presents additional design considerations when establishing blockchain architectures in heavily regulated cross-border markets, such as the derivatives market. In this case, a successfully compliant blockchain architecture will consider the additional accountability and transparency requirements imposed in U.S. regulations such as the FCRA, GLBA, Know Your Customer (KYC), and the Dodd-Frank Wall Street Reform and Consumer Protection Act (H.R. 4173, particularly Titles VI and VII). Some particular considerations for blockchain design in the US include the following requirements:

- Standardized derivatives must be traded on regulated exchanges or swap execution facilities to increase transparency of swap trading (H.R. 4173).
- Standardized derivatives must be moved into central clearinghouses (H.R. 4173).
- Swap dealers are subject to additional **record-keeping** and **reporting** requirements (H.R. 4173).
- The Financial Privacy Rule (15 U.S.C. §§ 6801–6809) requires financial institutions to provide consumers with a **privacy notice** (similar to transparency requirements under GDPR and CCPA).
- Financial Privacy Rule also gives consumers the right to **opt-out** of information being shared, which is arguably more protective than CCPA due to the former's emphasis on sharing vs. selling information. (Selling information is narrower in scope.)
- FCRA provisions which require creditors to **rectify** and **delete** information within 30 days of receipt of dispute (similar to GDPR window but shorter than CCPA).
- KYC core principles of transaction monitoring and customer identification for both banks and businesses (particularly including intermediaries).

4.1 Personal Data vs. Personal Information vs. Personally Identifiable Information

Sections 3.1 and 3.2 contrasted the GDPR's definition of personal data with the CCPA definition of personal information; however, when discussing these terms in the context of U.S. industry, it is important to include the more commonly-used concept of personally identifiable information (PII).

PII is "any information about an individual maintained by an agency, including (1) any information that can be used to distinguish or trace an individual's identity, such as name, social security number, date and place of birth, mother's maiden name, or biometric records; and (2) any other information that is linked or linkable to an individual, such as medical, educational, financial, and employment information."[37] The examples of PII provided by NIST cover essentially the same data types as those under the CCPA and GDPR:[38]

- Name, such as full name, maiden name, mother's maiden name, or alias
- Personal identification number, such as social security number (SSN), passport number, driver's license number, taxpayer identification number, patient identification number, and financial account or credit card number
- Address information, such as street address or email address
- Asset information, such as Internet Protocol (IP) or Media Access Control (MAC) address or other host-specific persistent static identifier that consistently links to a particular person or small, well-defined group of people
- Telephone numbers, including mobile, business, and personal numbers
- Personal characteristics, including photographic image (especially of face or other distinguishing characteristic), x-rays, fingerprints, or other biometric image or template data (e.g., retina scan, voice signature, facial geometry)
- Information identifying personally owned property, such as vehicle registration number or title number and related information
- Information about an individual that is linked or linkable to one of the above (e.g., date of birth, place of birth, race, religion, weight, activities, geographical indicators, employment information, medical information, education information, financial information).

When deciding which information is PII, NIST guidelines specify the concepts of *distinguishability, traceability,* and *linkability*. Linkability in particular is similar to the GDPR concepts of indirect and direct identification, in which data can be deemed PII/or personal data if it can be used to link back to an individual through another piece of personal information. (See Annex 2 for a more in-depth review of personal data under the GDPR). In PII as in the GDPR, linkable data could still be in separate but related databases. Data can be considered unlinked, however, if there are security controls that effectively segregate the information sources.[39] In both regulatory frameworks, it is best practice for organizations to document their logic in determining whether a particular data type is PII or not. It will also be important to remember this issue of linkability / indirect identification when evaluating the unique compliance solutions of SSI architectures such as Sovrin, which will be addressed in Section 6.

[37] NIST Special Publication 800-122, "Guide to Protecting the Confidentiality of Personally Identifiable Information" (April 2010), 2-1. Available at: https://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication800-122.pdf

[38] Ibid, 2-2.

[39] Ibid, 2-1.

Nonpublic Personal Information (NPI)

NPI, a term used in financial privacy regulations such as the GLB Act and FCRA, refers to "any 'personally identifiable financial information' that a financial institution collects about an individual in connection with providing a financial product or service, unless that information is otherwise 'publicly available.'"

NPI is:

- any information an individual gives you to get a financial product or service (for example, name, address, income, Social Security number, or other information on an application);
- any information you get about an individual from a transaction involving your financial product(s) or service(s) (for example, the fact that an individual is your consumer or customer, account numbers, payment history, loan or deposit balances, and credit or debit card purchases); or
- any information you get about an individual in connection with providing a financial product or service (for example, information from court records or from a consumer report).

Source: Federal Trade Commission

4.2 Personal Data Requirements for KYC-AML Compliance

KYC in the U.S. consists of a three-step process: 1) Customer Identification Program (CIP), which establishes a customer's identity; 2) Customer Due Diligence (CDD), which verifies the customer's identity; and 3) Enhanced Due Diligence (EDD), in cases where the CDD inquiry led to a high risk determination. Specific CDD and EDD procedures depend on the risk profile of the bank, but CIP (and basic CDD) is generally standard across financial institutions in the U.S.[40]

4.2.1 CIP

A Customer Identification Program (CIP) comprises part of a bank's BSA/AML compliance program, functioning primarily to ensure that the bank has a reasonable belief that it knows the true identity of each customer.[41] In order to comply with Section 326 of the USA PATRIOT Act, banks are required to collect and record certain personally identifiable information from customers as part of their CIP. Before allowing a customer to open an account, a bank must collect the following identifying information, *at minimum*:

- Name
- Date of birth
- Address (physical)

[40] See, generally, https://www.pwc.com/gx/en/financial-services/publications/assets/pwc-anti-money-laundering-2016.pdf.

[41] See https://www.ffiec.gov/bsa_aml_infobase/pages_manual/olm_011.htm

- Identification number[42]

4.2.2 CDD

Banks also need to verify that this information is correct. It can do so through either documentation or non-documentary methods. If a bank verifies through documentation, it must collect:[43]

- For an individual: unexpired government-issued identification evidencing nationality or residence and bearing a photograph or similar safeguard, such as a driver's license or passport.
- For a non-individual person (e.g. corporation, partnership, trust): documents showing the existence of the entity, such as certified articles of incorporation, a government-issued business license, a partnership agreement, or trust instrument

If a bank uses non-documentary verification, CIP-approved methods include: [44]

- contacting a customer
- independently verifying the customer's identity through the comparison of information provided by the customer with information obtained from a consumer reporting agency, public database, or other source
- checking references with other financial institutions
- obtaining a financial statement

4.2.3 Cross-border financial transactions

Section 312 of the USA PATRIOT ACT requires covered U.S. financial institutions to apply enhanced due diligence (EDD) when establishing or maintaining an account for a foreign bank that is operating:[45]

- under an offshore banking license
- in a jurisdiction found to be non-cooperative with international anti-money laundering principles; or
- in a jurisdiction found to be of primary money laundering concern under §311 of the USA PATRIOT Act.

Where EDD is required, organizations are required to take reasonable steps to:

[42] In the US, the "identification number" means a taxpayer identification number, or TIN (See 6109 of the Internal Revenue Code of 1986 [26 USC 6109]). For non-US persons, an identification number could be one or more of the following: TIN; a passport number and country of issuance; an alien identification card number; or a number and country of issuance of any other unexpired government-issued document evidencing nationality or residence and bearing a photograph or similar safeguard.

[43] Documentary verification methods are described by FinCEN in 31 CFR § 1020.220, paragraphs (a)(2)(ii)(A). See: https://www.ffiec.gov/bsa_aml_infobase/pages_manual/regulations/31CFR1020_220.pdf.

[44] 31 CFR § 1020.220 paragraph (a)(2)(ii)(B)(1)

[45] Information obtained from: PWC, *'Know Your Customer' Quick Reference Guide*, p. 317. Available at: https://www.pwc.com/gx/en/financial-services/publications/assets/pwc-anti-money-laundering-2016.pdf

- conduct appropriate enhanced scrutiny
- determine whether the foreign bank itself offers correspondent accounts to other foreign banks (i.e., nested accounts) and, as appropriate, identify such foreign bank customers and conduct additional due diligence on them; and
- identify the owners of such foreign bank, if its shares are not publicly traded

PWC's KYC guidance adds, "In circumstances where a financial institution establishes a relationship with a customer remotely, the institution will need to employ non-documentary methods to verify the identity of the client since it will not be able to use a document to compare the customer to the photo identification, or it will need to establish appropriate reliance agreements in order to rely on a third party who will conduct CIP on behalf of the institution."

4.2.4 Record-keeping

Banks must hold records of the former personal information (name, date of birth, address, and ID number) for five years after the account is closed.[46] Records should also contain descriptions of any document or non-documentary method(s) used in the verification process. Documentary descriptions should include any identification number contained in the document, the place of issuance and, if any, the date of issuance and expiration date.[47]

4.2.5 Suspicious Activity Reporting (SAR)

Banks are required to have an SAR system in place as part of their BSA/AML compliance regime. Effective SAR systems have five key components:

- Identification or alert of unusual activity (which may include: employee identification, law enforcement inquiries, other referrals, and transaction and surveillance monitoring system output)
- Managing alerts
- SAR decision-making
- SAR completion and filing
- Monitoring and SAR filing on continuing activity

Part of the SAR involves monitoring customer transactions, although the sophistication and depth of the monitoring system depends on the bank's risk profile. CIP, described above, is a part of an effective monitoring system, along with transaction-based (manual) and surveillance (automated) systems.[48]

Surveillance systems are important to describe in a little more detail due to the GDPR's heightened requirements for automated decision-making (Article 22). SAR surveillance monitoring systems

[46] 31 CFR § 1020.220 paragraph (a)(3)(i)(A).

[47] 31 CFR § 1020.220 paragraph (a)(3)(i)(B).

[48] https://www.ffiec.gov/bsa_aml_infobase/pages_manual/OLM_015.htm

identify individual transactions, patterns of unusual activity, or deviations from expected activity through either rule-based or intelligent monitoring. As the FFIEC describes,

> "Intelligent systems are adaptive and can filter transactions, based on historical account activity or compare customer activity against a pre-established peer group or other relevant data. Intelligent systems review transactions in context with other transactions and the customer profile. In doing so, these systems increase their information database on the customer, account type, category, or business, as more transactions and data are stored in the system."

The important point here is that intelligent systems develop customer-based profiles based on a multiple data sources around demographic and other characteristics, rather than just solely on transactional activities.

4.2.6 Informational Notices and Opt Out

Financial institutions falling under the scope of Subtitle A of Title V of the GLB Act and handling nonpublic personal information need to meet specific informational disclosure requirements.[49]

- Initial notice: provide clear and conspicuous notice to the customer or consumer that accurately reflects privacy policies and practices. Information should be disclosed before or upon entering into a customer relationship, and prior to disclosing any personal information to a third party.[50]
    - Notice cannot be given verbally (unlike the GDPR).[51] This is because it is important that the customer be able to keep a record of the notice and policies.[52]
- The notice should contain:[53]
    - The categories of nonpublic personal information collected.
    - The categories of nonpublic personal information disclosed to others.
    - The categories of affiliates and non-affiliated third parties to whom you disclose nonpublic personal information.[54]
    - A separate statement of the categories of personal information disclosed to third parties (if applicable) and the categories of any contracted third parties.
    - An explanation of the consumer's right to opt out of the disclosure of nonpublic personal information to nonaffiliated third parties, including the methods for doing so.
    - Any disclosures required under section 603(d)(2)(A)(iii) of the Fair Credit Reporting Act.

[49] For a list of entities that fall out of scope, see 12 CFR § 1016.1(b), accessible at: https://www.law.cornell.edu/cfr/text/12/1016.1

[50] 12 CFR § 1016.4, see (b) for exemptions

[51] 12 CFR § 1016.9 (d)

[52] 12 CFR § 1016.9 (e)

[53] These categories can be found under 12 CFR § 1016.6 (a)

[54] Exceptions can be found in 12 CFR § 1016.14 and § 1016.15. https://www.law.cornell.edu/cfr/text/12/1016.14

- Policies and practices with respect to protecting the confidentiality and security of nonpublic personal information, including (in general terms) who is authorized to access the information.[55]
- The notice should be given annually.[56]

> **Opt out** means "a direction by the consumer that you do not disclose nonpublic personal information about that customer to a nonaffiliated third party, other than as permitted by §§1016.13, 1016.14, and 1016.15."
>
> - *12 CFR § 1016.10(a)(2)*

4.3 Contrasting GDPR/CCPA with KYC-AML

When contrasting KYC-AML requirements with data privacy legislation, it is important to re-emphasize that the CCPA cannot conflict with federal regulations regarding financial institutions (i.e. Section 4.2) such as the GLB Act or the FCRA (see CCPA 1798.45[d][e]).

Figure 3: Informational Notice Obligations

Disclosure Requirement	GDPR	CCPA	KYC-AML
Categories of PD collected	Yes	Yes	Yes
Specific pieces of PD/PI collected	No (unless under access/portability right)	Yes	No
Categories of PD disclosed to third parties	Yes	Yes	Yes
Categories of third-parties to whom PD is disclosed	Yes	Yes	Yes
Explanation of individual right to opt-out	Yes (if applicable)	No – only able to opt out ("object") of sale of PI to third-parties	Yes
Explanation of purpose and legal basis for processing	Yes	Must explain business/commercial purpose. Not obliged to explain legal basis	NA (typically implicit in the context of the relationship)

[55] Note: you do not have to disclose any technical information about the safeguards used. 12 CFR § 1016.6 (c)(6)(ii).
[56] 12 CFR § 1016.5

Explanation of lawful basis of processing	Yes	No	NA (typically implicit)
Notice of cross-border transfers	Yes	No	No
Retention times	Yes (or criteria used to determine)	No	No, although could interpret via §1016.6(a)(8)
Data subject (or consumer) rights	Yes	Yes (but different rights)	Yes (but only opt-out)
Provide easy access to exercising rights	Yes	Yes – at least two means including toll-free number and email	Yes
Describe security and confidentiality measures	Yes	No	Yes

Figure 4: General Obligations/Requirements

Obligation	GDPR	CCPA	KYC-AML
Information	Yes (different requirements)	Yes (different requirements)	Yes (different requirements)
Access	Yes (different requirements)	Yes (different requirements)	No, although arguable under transparency requirements
Deletion	Yes	Yes	Yes (FCRA)
Rectification	Yes	Not in CCPA	Yes (FCRA)
Portability	Yes	Yes	No
Objection	Yes	Yes (only to sale of PI)	Yes (aka opt-out of disclosure)
Nondiscrimination based on privacy preferences	Somewhat (only re: consent)	Yes	No
Record-keeping	Yes (different requirements)	Yes (different requirements)	Yes (different requirements)
Data Retention	Allowed until purpose has been fulfilled or there is no longer a lawful basis	Not in CCPA	Required to retain CIP info 5 years after end of relationship
Perform DPIAs	Yes (for new, potentially risky technologies that process PD)	Not in CCPA	NA, although similar risk assessment procedures required (e.g. under GLBA 501(b))
Hire/Appoint a DPO	Yes (if regular, large-scale processing of special category data)	No DPO role exists in CCPA	No
Hire/Appoint Representative if not established in territory	Yes (if large-scale processing of special category data and no establishment in EU	No, businesses don't need separate representative in California in any situation	NA
Explain automated decisions	Yes, if processing is regular and could legally or similarly significantly affect DS	No	No
Response time to consumer/subject requests	Within 30 days	Within 45 days	NA

Protection of "minors"	Parental consent for processing of PD required under 16 years old	Parental consent for sale of PI required under 16 years old	NA
Security	Yes (Art 32)	No, although arguably implied in 1798.150 (a)(1).	Yes (under FTC Safeguard Rule)
Data Protection by Design and Default	Yes	Not in CCPA. Arguably yes under FIPPs.	Arguably implicit in security and access control requirements
Data breach communication	Yes	Not in CCPA but regulated in state laws (e.g. California Civil Code s. 1798).	Yes: GLBA Section 501(b); 12 CFR 364, Appendix B
Cross-border Transfers	Strict requirements for transferring outside EEA (Arts 44-50, see Annex)	Not in CCPA	Yes, see Section 4.2.3

It is also important to note the interaction between different regulatory obligations. In the authors' opinion, the GDPR and KYC-AML have no directly contradictory requirements. In instances where there are differences in regulatory requirements (see above), the GDPR gives leeway to companies that need to comply with specific data retention and non-disclosure requirements in other national laws. For example, in cases like the GLBA requirement to hold CIP data for five years after the customer relationship has ended, the financial institution would be justified in this retention limit under the "legal obligation" lawful basis for processing (GDPR Article 6(1)(c)). Other potentially conflicting obligations popularly discussed under the GDPR, such as erasure (Art. 17), would most not prove contradictory to financial institutions' KYC-AML obligations for the same reasons of legal obligation.[57]

However, some obligations in one regulation can force additional compliance obligations in another regulation. For example, if a business has to comply with both KYC and the GDPR, KYC requirements will likely require the business to comply with stricter GDPR requirements than it might have to otherwise. To explain through an example: central to best practice in KYC risk management is the concept of transaction monitoring, which involves the establishment of expected user behavior for banks or intermediaries to check for anomalous account activity. In order to create a successful expected user behavior, the entity typically must perform individual participant tracking of transactions and content in order to establish a profile by which to understand what "expected" means for that user. If this were to occur to an EU data subject, for instance, additional requirements would be imposed on a business due to the fact that these

[57] In other words, a data subject's grounds for requesting erasure – established under Article 17(1) – are not valid in these circumstances.

profiling processes are often automated and have legal or similarly significant impacts on the entity being monitored. This would likely require the entity to fall under the scope of additional GDPR requirements such as the need to hire a Data Protection Officer (DPO), establish an EU Representative (if the business has no establishment in the EU), and document Data Protection Impact Assessments (DPIAs) for these automated processing technologies. These requirements are only mandated for businesses that reach this threshold of regular and systematic large-scale monitoring.[58] Regardless, the financial institution would also need to explain the logic and envisaged consequences of any automated decision-making processes under GDPR Article 13(2)(f) and 14(2)(g).

Another common practice in CDD is the verification of CIP information through photo-identification. The information gleaned about a person from a photograph – e.g. race, sex, perhaps ethnicity – is classified under the GDPR as "special category" personal data. The processing of special categories of data come with additional obligations and prohibitions (see Annex 1 for a detailed overview). While financial institutions would typically have a legitimate lawful basis of processing (i.e. legal obligation under Article 6 and 9[2]), these institutions would most certainly have to keep a record of their processing activities (regardless of business size), and most likely have to hire a DPO and perform DPIAs if the processing is regular and on a large scale. However, whether these requirements were triggered would depend on how the photo ID verification was handled. For example, if the information was physically disclosed by the data subject to an authorized party within the institution without scanning or otherwise recording the photograph, then one could potentially argue that the personal data was not "collected" by the financial institution. As described in Section 4.2.4, CDD record-keeping under KYC does not require verification documents to be recorded; instead, the method of verification itself must be logged. Thus, we would recommend either using an alternative form of verification such that it does not collect special category data (at least for people in the EU) or using photo verification through the use of a Self-Sovereign Identity (SSI) architecture that would allow only need-to-know data to be disclosed and recorded. Technical implementations for achieving this will be discussed in Section 6.

[58] See generally Arts. 27 (EU Representative), 35 (DPIAs), 37 (DPOs), and 22 (regarding restrictions on automated profiling).

5 Fitting Blockchain into Legal Scope, Roles, and Definitions

Legal roles are difficult to assign in a blockchain network due to the distributed and often ambiguous control and decision-making power of nodes on a network and even network operators themselves. Because control is dependent on the technical configuration of the network, it is impossible to assign fixed legal roles (e.g. data controller, processor) to blockchain actors (e.g. node, network operator).

5.1 Blockchain Architectures in Brief

The task of attributing the regulatory requirements of conventional computer network roles to agents within a blockchain network is confounded by the variety of blockchain networks, which vary in the types of roles they establish, as well as the responsibilities for the same role.

Public blockchains are highly transparent and fully decentralized ecosystems that achieve the special characteristic of being 'trustless'. Yet they are also notoriously slow, resource intensive, and ill-suited for adherence to the KYC and due diligence requirements to which enterprise systems and financial applications are beholden. *Private blockchains* address these issues by embracing centralization, but this necessarily limits the openness and potential for broad adoption of such networks. Because the nodes maintaining a private ledger are exclusively controlled by the network operator(s), private blockchains allow said operators the rights to override, edit or delete entries on the blockchain as required. While this helps network operators comply with certain legal requirements (e.g. FCRA, GDPR, and CCPA provisions of rectification/deletion), they reintroduce many of the very issues of intermediary reliance that blockchains were intended to address. Further, such networks arguably achieve less transparency and accountability (to the users of such a network) by design than their public counterparts because only entities approved by the network operator may maintain a record of the ledger.

Permissioned ledgers have been proposed as a compromise between public and private that allows for a high degree of customizability, such that a balance (as required for a given blockchain-based community, product, or utility) can be struck between the two ends of a spectrum that public and private ledgers represent. A defining characteristic of permissioned ledgers is that they programmatically establish allowed behaviors and levels of data access for each category of participant - significant for achieving 'privacy by design'.

Metadata in Blockchain vs. TCP/IP

Blockchain is unique from existing Internet architecture because personal data does not *intrinsically* need to be stored on-chain. To understand exactly why personal data need not intrinsically be stored on-chain, consider the difference between the contents of a transaction. In a TCP/IP system, every interaction intrinsically encodes personal data in the form of an IP address. IP addresses are explicitly referred to as examples of personal data in GDPR Article 4(1) and CCPA 1798.140(o)(1)(A). Yet a blockchain transaction can be made between parties that have no knowledge of each other's identity. The tool behind this is called a decentralized identifier, or DID (which will be described in Section 6). Essentially, a unique identifier is created for each party upon the establishment of a relationship, which cannot be linked back to an original 'IP address' equivalent or used to track a single party across multiple relationships (because the DID will be different for each one). As a result, the anonymity possibilities afforded by the 'metadata' of a blockchain network far surpass those embedded in the transport layer of the Internet.

5.2 Network Operators

A perfectly decentralized blockchain network arguably places the network operator in a technical role analogous to a traditional ISP or telecommunications company. The network operator essentially provides the connection and interaction protocol for users (nodes) and facilitates interactions between users. The network operator does not necessarily need to access the data content, and since source and destination metadata on a blockchain can be configured to be less personally identifiable than an IP address, the network operator may not necessarily "process" personal data at all.

However, this network design may not comply with certain sector-contingent regulatory requirements. Markets that are subject to KYC, for example, require intermediaries or "middle men" to monitor the transactions on their network for security and transparency purposes. If the network operator was determined to be an intermediary of this nature, then the network operator would have to be able to attribute transactions to particular parties, which would then require the metadata to contain personal information/data under both GDPR and CCPA definitions. As a result, the network operator would be subjected to the scope of the GDPR and CCPA (assuming the operator met the other scope requirements). If the network operator runs a blockchain network facilitating a derivatives market under the scope of U.S. federal law, then the network operator may reasonably be placed in the role of a central clearinghouse or swap execution facility. In this case, it would have to monitor and record additional information about the entities engaging in transactions on its network, as well as information regarding contents of those transactions (as per H.R. 4173).

In order to achieve regulatory compliance in its monitoring and recording practices, the network operator would have to design a system centralized enough to allow the operator to do so, but permissioned enough to allow for consumer privacy requirements regarding rectification, deletion, and opt-out under the GDPR, CCPA, GLBA, and/or FCRA. This would most likely require the network operator to establish a permissioned ledger whereby transactions between nodes could occur on private ledgers monitored and accessible only by the network operator, with hashes of those transactions recorded on the public blockchain. The hashed representations of the transactions on the public chain would provide transparency and accountability *of the network operator* to regulators, who would be able to know when a transaction was edited or changed, and then be able to investigate the nature of the changes.[59] Section 6 will describe in further detail how a type of permissioned ledger can function to meet regulatory requirements.

Thus, at least in the case of derivatives markets, the network operator is most likely forced into the position of a data controller if it must comply with GDPR, KYC and other aforementioned laws. In order to limit liability and onus, it could potentially outsource its monitoring or record-keeping requirements to a third-party auditor.[60] In this case, the auditor or any subcontractors would probably be data processors of the network operator. However, in CCPA terms, it is unlikely that these auditors would need to be compliant due to their restricted decision-making power over any personal data.

[59] This is because a smart contract transaction could be regulated such that each step on the process of data adaptation (request, verification, alteration, confirmation) could represent a separate transaction. Each step can only be fulfilled by certain parties (e.g. request for alteration can only be fulfilled by the data subject; verification and alteration can only be executed by the network operator; confirmation can only be fulfilled by all parties).

[60] E.g. the network operator facilitates the transactions but restricts its access to the record itself, instead giving that permission to the auditing entity. This is called ‘audit-enabled privacy’, which is particularly useful in financial transaction use-cases and is already being implemented (e.g. through Hyperledger Fabric). It will be further discussed in Section 6.

Is a Hash Pseudonymous or Anonymous Data?

Hashes are one-way cryptographic functions commonly used to de-identify personal data. NIST guidelines classify hashed data as pseudonymous PII with a confidentiality impact of *low* (SP 800-122, p.4-5). NIST also notes that some regulations (such as HIPAA) do not allow for de-identification techniques where the code is derived from the PII itself (45 CFR §164.54 (c)(1)). The EU Article 29 Working Party also classifies hashing as a pseudonymization technique rather than anonymization (Opinion 05/2014 on Anonymization Techniques, p. 20).

The reason it is referred to as pseudonymous rather than anonymous information is because the correct value for a particular record can be derived through the hash if the range of input values are known. However, some techniques, such as a keyed-hash function with a stored key, make common vulnerabilities (e.g. brute force attacks) computationally impractical. It is uncertain whether NIST or the Article 29 Working Party would view the hashing techniques in permissioned and ledgers as sufficiently effective enough to render the data effectively anonymous to parties that only have access to the hash itself. Indeed, as the Article 29 Working Party notes, the effectiveness and classification of the technique as pseudonymous or anonymous needs to be evaluated on a case-by-case basis (WP216, p.23).

In this Report, we assume that hashed data is no longer personal data to parties without access to the input values. We believe this is a fair assumption, as long as the private or permissioned ledgers (or off-chain datasets) with the input values are constructed with appropriate access controls and safeguards. However, it is important to note that this assumption has not received clarity by the Article 29 Working Party.

5.3 Nodes

If we assume that nodes process transactions containing personal data, then each node processing the transaction will be required to comply with the GDPR/CCPA, as long as the nodes/personal data fulfill the other scope requirements outlined in Section 3.[61] However, depending on the structure of the blockchain architecture, the fact that one node falls under the scope of the GDPR/CCPA could cause all other nodes on the network to fall under the scope of the law.

5.3.1 Scenario 1: Personal data is published on-chain

In a decentralized public ledger (and possibly in certain permissioned ledger designs), if personal data of an EU resident is processed by a node and published on-chain, the entire ledger and all nodes will fall under the scope of the GDPR. This is due the fact that all nodes will maintain a

[61] This means that the personal data would have to refer to a California resident (to fall under CCPA scope); in the case of the GDPR, if the personal data referred to a person in the EU (commonly interpreted as an EU resident), then it would force the business into GDPR compliance as long as the processing were for commercial or monitoring purposes.

record of the ledger, which amounts to "processing" under Article 4(2). Similarly, if personal data of a California resident is published on-chain by a node (and either the node or the resident are in California during the time of transaction and the node meets the threshold of one of the three bullets that define a "business"), then the California resident's personal data falls under the scope of the CCPA and all nodes must be compliant.

This architecture-driven snowball effect is clearly undesirable for most nodes, and it also presents further GDPR compliance problems for nodes and regulators alike. Perhaps the most poignant of these issues is international transfers. If any node stores their ledger outside the EEA or a country with EU-determined legal "adequacy", then that node must inform all other nodes of the geographic location of the data processing and further set up contractual clauses *with each node* to legally bind itself to GDPR compliance.[62]

Recommendation: These dual issues (snowball effect and cross-border transfer obligations) can be resolved through the use of private or permissioned ledgers. For example, an administrator can assign nodes different restrictions (or 'permissions') and allow nodes to perform different types of transactions. In such a scenario, certain nodes could be given permission to process personal data of EU or California residents, as long as they are either in the EU or the network administrator verifies that appropriate legal safeguards are already in place to comply with the relevant regulation. Then, once the 'transaction' is completed, the node publishes a hashed record of the transaction to the ledger.[63] The ledger would remain free of personal data and thus free of GDPR/CCPA compliance, but the node(s) engaging in the transaction would have to be compliant.[64]

[62] The different forms of contractual clauses/agreements can be found in Chapter 5 of the GDPR, particularly Articles 45-47. Although, depending on the nature of the contract between the controller the data subject, the controller and all other nodes may be able to circumvent contractual clauses and instead justify the transfer under a derogation clause in Article 49(b) or (c), arguing that the transfer is necessary to either perform or conclude the contract with the data subject. However, a controller will have to justify that the transfer itself was necessary, which is a difficult argument to make considering the network could have been permissioned to restrict third-country transfers.

[63] A hash refers to a one-way (i.e. irreversible) cryptographic function that allows parties to verify that a set of data is the same without sharing the underlying data. On a blockchain – particularly a public one – it provides a mutually established receipt of a transaction recorded on the immutable ledger, which makes it useful for both dispute resolution and authentication over the contents of a set of data without disclosing the data to other parties on the chain.

[64] Again, this assumes that the node meets one of the additional thresholds for CCPA compliance.

KYC and on-chain personal data

Personal data (or NPI) does not need to be held on-chain in order to meet KYC obligations of CIP, CDD, and record-keeping. If a financial institution wishes to use a blockchain ledger to hold personal information, it can be held on a private or permissioned ledger accessible only to parties that meet rigorous access controls. The information (or its location off-chain) could then be referenced through an encrypted pointer held on the public chain. Alternatively, the information could be hashed, and the hash itself could be located on-chain. This would be a more secure method of referencing personal data on-chain. In either case, auditors or regulators could either be given permission to access the ledger directly, or alternatively they could be given a decryption key or (in the case of a has) the underlying data and verify by viewing the data or matching the hash results.

Another important aspect of KYC-AML is customer transaction monitoring. One might worry that if node transactions occurred on private ledgers and with pseudonymous identifiers (see DIDs in Section 6), financial institutions would not be able to effectively monitor transactions. In this case, we would recommend setting up a permissioned ledger for a customer transaction system, whereby either the financial institution or a subcontracted auditor was given limited access to transactional ledgers that are off the public chain.

5.3.2 Scenario 2: No personal data on-chain

Now let us assume that no personal data is processed on-chain, but instead a node on the network is an establishment located in the EU or California.[65] Under the GDPR, if the node located in the EU processes personal data as a part of its business activities (e.g. off-chain), then the entity must comply with the GDPR, regardless of any affiliation to blockchain. Yet since the public ledger holds no personal data, the network and its operations remain outside the scope of both the GDPR and the CCPA.

5.3.3 Are Nodes Data Controllers or Processors?

5.3.3.1 Scenario 1: Node processes personal data through the ledger

[65] At this point it becomes a somewhat ambiguous GDPR question regarding whether a node located in the EU would qualify as an "establishment" representing that business. For example, If I work for Company X based in the U.S., but am on business in Belgium and happen to process a transaction as a node on a blockchain network during my trip, do I fall under the scope of the GDPR? While the answer to this scenario is probably negative, what if I instead worked remotely in Belgium from a home office and consistently acted as a node on a blockchain network on behalf of Company X? In that case, my activities may indeed cause Company X to meet the Recital 22 threshold of "real and effective exercise of activity through stable arrangements" or the *Weltimmo* threshold of a "single representative."

In the above scenario in which personal data is published to the ledger, it becomes legally ambiguous whether each node processing transaction via the ledger (rather than performing the transaction itself) is a data controller or a data processor. One could argue that each node is effectively a joint controller because by recording the transaction on its ledger, it holds decision-making power by attesting to the existence – or truth – of the transaction.[66] Yet one could also argue that each node has little control over the data, as it has no say regarding how the personal data will be processed, stored, or to whom it will be disclosed. In this light, the node acts more as a data processor that performs an operation on personal data on behalf of the "controller," or the node that published the transaction to the ledger. Operating under this latter framework, these nodes would fall out of the scope of the CCPA due to their lack of decision-making power. In any case, the definitional ambiguity creates large uncertainty regarding the compliance obligations of a node processing merely the ledger and not the transaction itself.

Yet this ambiguity can be resolved in a permissioned ledger where the network defines parameters such as decision-making power, international transfers, and access to personal data. Since a permissioned ledger – if deployed correctly – would not require personal data to be directly published on-chain in order to achieve accountability/transparency requirements, these compliance dilemmas would not arise.

5.3.3.2 Scenario 2: Intra-transactional personal data processing

Consider the following scenario: two entities in the EU process personal data on a private ledger as part of a transaction on a derivative market. This transaction is then hashed and recorded on the public blockchain. Since the contract contains both parties' personal data, both parties are therefore data subjects, and the network operator (brokering and recording the transaction for regulatory accountability) becomes a data controller. Yet since it takes both parties to mutually attest to the transaction in order for it to be written to the blockchain, both parties also arguably become "joint controllers" of each other's personal data.[67] Note also that all other nodes on the network affirming the transaction and writing it in the ledger would not fall under the scope of the GDPR since they only have access to the hash of the contract.

The CCPA offers a more straightforward puzzle since it omits data processors from the scope of compliance, and instead refers to a "business" and a "consumer" as its roles. Yet even these binary distinctions present problems when determining who exactly is the "business" in certain scenarios. In the majority of cases, it is likely that the node that liaises the transaction with the California resident will be the "business" (recall the earlier scenario of a node processing a Californian's personal data in a permissioned ledger). However, the "business" could also arguably be the

[66] To explain this point: whether a block is written to the blockchain depends on the majority consensus of the nodes in the network. Each node, therefore, gets a "vote" whether to process that block or not. If the majority decides that it is an invalid (or "bad") block, then the block becomes abandoned (or "orphaned") by the other miners. This creates what is known as a "fork" in the blockchain, whereby the old route (containing the bad block) is essentially deleted since every new block operates on the premise that the invalid block was never written. This, in short, is a type of decision-making power that each node has over the personal data processed within a transaction.

[67] See Article 26 for reference to "joint controllers". Note: this is an oversimplified description of how a contract might be made. Depending on the blockchain protocol, each entity could individually attest to the contract without the other's consent (i.e. they would be attesting to their own version of the contract), and that self-prescribed "transaction" would be recorded on-chain without the other party's involvement.

network operator that determines the permissions and protocol of the network itself (e.g. Sovrin).[68] This would depend on the exact permissions of the network administrator, i.e. the amount of decision-making power they have over the personal data.

5.4 Conclusion: Leveraging Ambiguity

GDPR and CCPA were conceived with conventional network architectures and data infrastructures, not blockchain networks, in mind. Accordingly, the roles and corresponding regulatory obligations defined in both do not map well to the roles in a blockchain network. These ambiguities in role assignment make it exceedingly difficult for entities in a blockchain construct (as well as regulators) to know who needs to comply with which part of the regulation, and how they must do so if they indeed fall under the scope of the law. Nevertheless, we believe that discrete role assignment and regulatory compliance are both achievable. Businesses can view the ambiguity not as a hindrance but as an opportunity to define legal roles and compliance obligations. With knowledge of the defining attributes of a data controller, processor, business, and consumer, networks and nodes can be constrained by permissions in the network architecture to befit a particular role. With knowledge of the scope of data protection regulations, network architectures can be designed to remove nodes, groups of nodes, or even entire networks from GDPR or CCPA compliance.

[68] Recall from 1798.140(c)(1), a business needs to comply with the CCPA if it collects California residents' personal information **or determines the purposes and means of processing that personal information.** This caveat (bolded) could arguably cause the network administrator in a permissioned blockchain to fall under the scope of the CCPA, since after all it is the administrator, not the node, that defines the purposes and means of processing with reference to the blockchain.

6 Sovrin as a Use Case for Achieving Regulatory Compliance

6.1 Purpose and Structure

This section demonstrates how regulatory compliance can be achieved with distributed ledger networks. We consider distributed ledger models that we believe represent industry best practice in the field of regulatory compliance. We use one industry leader in particular – the Sovrin identity management system – as a case study to illustrate how distributed ledger technologies can help achieve regulatory compliance. By focusing on Sovrin, we do not mean to imply that Sovrin is the only, or the best, method for achieving compliance with GDPR/CCPA and KYC-AML regulations. Rather, due to limitations in scope, we chose to provide greater detail in the technical and legal implementation of a single model rather than giving a more superficial overview of many models. That said, alternative models and techniques – such as Hyperledger Fabric and Trusted Execution Environments – are briefly discussed as an addendum to the case study.

Given the unique nature of Sovrin's approach to digital identity, we begin with a primer on Sovrin's conceptual framework for digital identity. This primer serves two purposes:

- to provide a foundation for understanding how Sovrin helps meet GDPR/CCPA obligations
- to describe a technical breakthrough in how roles can be assigned to actors in digital systems that protects the privacy of individuals and entities

The primer is structured as follows:

- Brief introduction to the Sovrin protocol
- Description of the three dimensions across which Sovrin separates digital identity
- Discussion of what takes place at the intersections of these three dimensions
- Overview of the significance of the Sovrin approach

Following the Sovrin primer, we discuss how Sovrin achieves compliance with core data protection principles and privacy by design requirements specified by the GDPR and CCPA in Sections 2-4 of this Report. Section 6.3 evaluates Sovrin's compatibility (and, generally, a Sovrin-type approach) with core data protection principles under the GDPR. Section 6.4 discusses Sovrin's ability to meet data subject rights requirements under the GDPR. Note: this section focuses heavily on the GDPR and less so on the CCPA. This is because as demonstrated in Sections 2.3 and 4.3, CCPA requirements, while slightly different in scope and legal implementation, are technically the same if not weaker to implement than GDPR requirements. In other words, if Sovrin can support data subject rights under the GDPR, it can also support consumer rights under the CCPA.

6.2 Sovrin Primer

The Sovrin Protocol is a set of standards, patterns, and tools that facilitate trusted interactions. Sovrin is rooted in the principle of Self-Sovereign Identity (SSI). SSI places the individual or entity (called an Identity Owner) in charge of their own identity management, as opposed to the traditional approach that uses centralized trusted-third-parties to manage identities. In theory, this decentralization decreases the need for individuals to rely on third parties to manage their personal data and further decreases the risks and obligations placed on third parties for handling and storing personal data (or PII) themselves. Moreover, SSI would reduce the inherent data siloing that follows from centralized institutional management. In an SSI model, individuals have full autonomy to create their own identity, manage keys, and share personal data. SSI models such as Sovrin's are uniquely designed to achieve compliance for regulations such as GDPR and CCPA.[69]

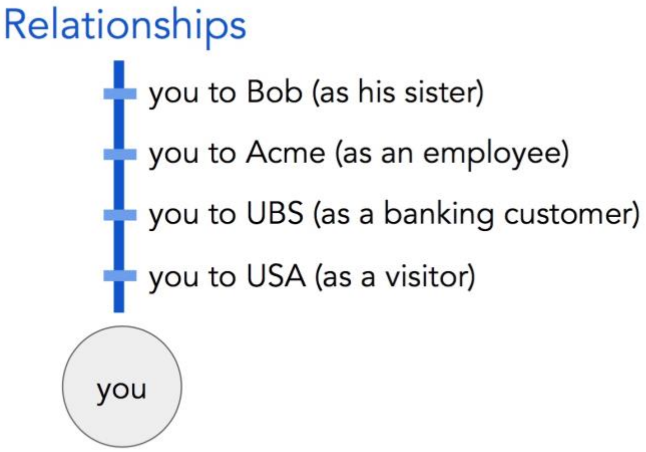
6.2.1 The Three Dimensions of Sovrin's SSI Model

A defining characteristic of Sovrin's approach to SSI is the separation of digital identity into three orthogonal dimensions:

(1) **Relationships** - who I am is contextual
(2) **Attributes** - what I can share and prove
(3) **Agents** - what or who can act on my behalf

6.2.1.1 Relationships

On the Internet, I find web pages and access services by contacting a domain name, which functions as a public reference to an IP address. In Sovrin, the domain name equivalent is called a 'public' decentralized identifier, or DID. A DID is 'decentralized' because there is no central authority or DNS equivalent that requires me to register my domain. Instead, if I want other entities on the network to be able to contact me, I publish my public DID to Sovrin's public ledger. Entities can choose to contact me in a similar fashion to typing in the corresponding URL. If I, a customer, contact a bank through its public DID, Sovrin established a private communication channel for me to transact with the bank and gives us each new private DIDs for that off-ledger relationship. At this point, even though we are in a relationship, the bank does not know anything about me because my DID has been uniquely generated for the relationship.


Relationships
you to Bob (as his sister)
you to Acme (as an employee)
you to UBS (as a banking customer)
you to USA (as a visitor)
you

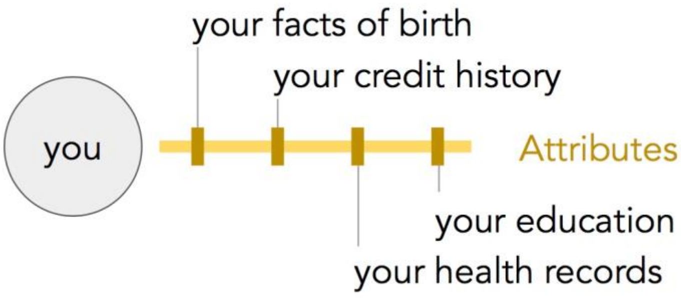
At this point, the TCP/IP analogy breaks down. On the Internet I would have to give the bank my IP address, which can be used to geo-locate me, track me across the web, or glean information about me from other companies that already have data about how my IP address interacts on the web. Yet in Sovrin, identifiers are not replicated across relationships. This prohibits unwanted or

[69] Sovrin is not the only SSI model. Others include uPort, SelfKey, and MyData. While SSI models in general are uniquely suited for achieving compliance, these latter models are less robust in their feature-offerings and completeness -- from a technical and design perspective -- than Sovrin.

cross-context correlations that lead to privacy loss and data leakage via cookies, web beacons, and other technologies on the Internet.

Note also that it is for exactly this issue of linkability that proxy - or indirect - attributes are covered under the GDPR's definition of 'personal data', the CCPA's definition of personal information, and the NIST framework's definition of PII. As a result, the scope of these regulations cover vast swaths of data and metadata in interactions over the Internet (e.g. cookies, web beacons, IP addresses, and any other information linked to them). Yet because identifiers cannot be correlated to the same Identity Owner unless the Identity Owner authorizes the link through a Credential, these types of information (cookies etc.) would arguably fall out of the scope of personal data/information. This would not only increase consumer privacy but also decrease liability for companies handling this data. Most of all, it would reduce the pervasive issue of proxy identification in a traditional TCP/IP architecture.

6.2.1.2 Attributes

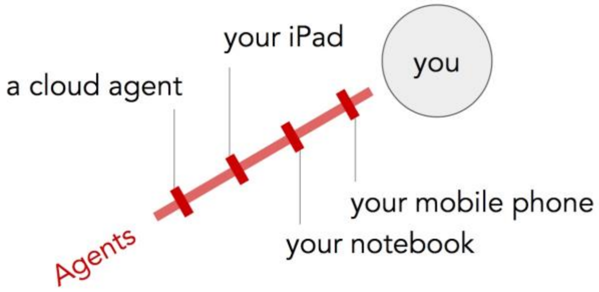


While I will have many DIDs over the course of my interactions on Sovrin, I only have one identity behind each DID. When I establish my identity on the Sovrin network, I am given a private ledger -- called a "Wallet" -- where I can securely store my attributes. My attributes are established through Claims, which are attestations made either by myself (the Identity Owner) or by another party. An Identity Owner who makes a Claim about another Identity Owner is called an Issuer. Claims are digitally signed by Issuers so that anyone receiving the Claim can know who issued it.

For example, say I have a Claim in my Wallet that represents my driver's license. In order for my driver's license to be valid, it had to be signed by my state's Department of Transportation. Upon producing the Claim for my driver's license, the DoT attested to multiple attributes associated with my driver's license (e.g. eye color, date of birth, height, etc.). Now, perhaps I want to establish a relationship with a bank through Sovrin, but the bank needs to verify my birthdate before I can open a bank account. Luckily, this attribute has been attested through the Claim representing my driver's license. I can now produce my birthdate along with the DoT's Claim that attests this particular attribute. (Note that I do not need to produce my entire driver's license - just the attribute necessary.) The bank will now be able to see that my DID now has a "date of birth" attribute linked to it, attested to by the DoT. If the bank agrees that this is a trusted Issuer, then it will proceed and open an account for me. If the bank needs more information, the process repeats.

Before moving on, it is important to note how DIDs work with attributes in situations like this. When the DoT issued a Claim about my attributes, it made a Claim to the DID associated with me. Yet this Claim is not intrinsically linked to the DID such that I can only use it if I am using the DID. Instead, the Claim was stored on my private ledger (Wallet) so that I could reproduce it for other pairwise relationships in which I have a different DID than the one in which the Claim was originally produced. This gives me added privacy because Claims cannot be linked together unless I choose to do so for that particular relationship. (For example, if the bank needed to know my birthdate and blood type, I could produce the DoT's Claim about my birthdate and my doctor's Claim about my blood type. These Claims would now be linked to the same private DID used for my relationship with the bank.)

6.2.1.3 Agents

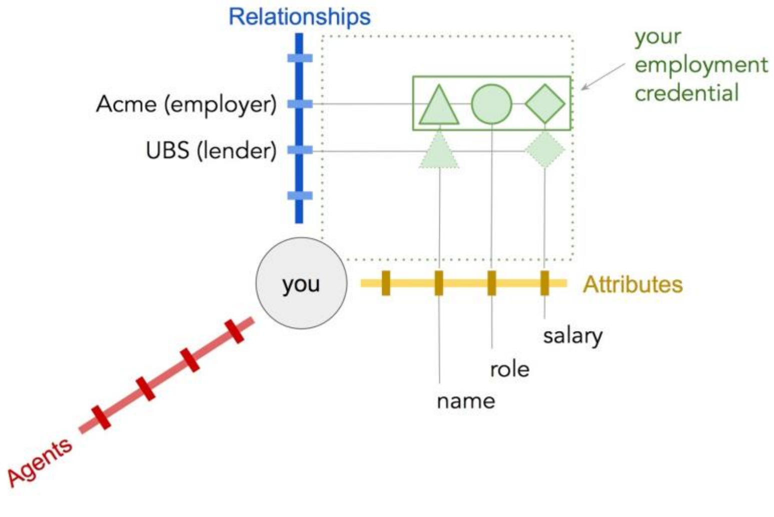


Agents are devices or services that fulfill tasks for us. Like in a traditional Internet network, devices are linked to our identities and also to the network through IP addresses. These can include a mobile phone, a physical notebook, software, or a third-party service provider. On a Sovrin network, I can only have one agent per device. I can also give an agent an endpoint, which can then serve as a simple message proxy, or I can even have it interact with other entities outside the Sovrin network.

6.2.2 Interactions Between Dimensions

6.2.2.1 The Relationship-Attribute Plane

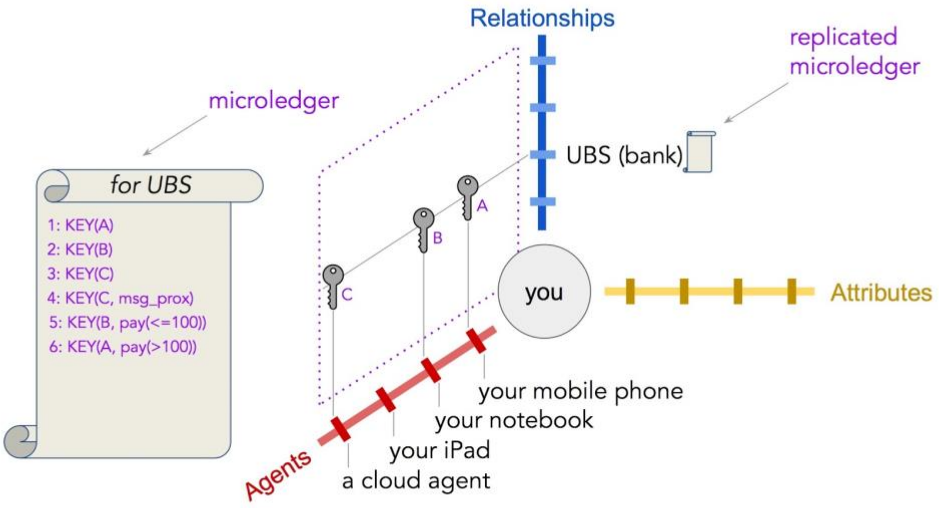
For every relationship, I can choose which attributes about my identity I wish to share and 'link' together as belonging to me via my pairwise DID. As referenced in an earlier example, it is common for entities to need more than one attribute over the course of a relationship. For example, in KYC-AML processes for CIP, financial institutions need at minimum four attributes: name, date of birth, address, and identification number. Through Sovrin, I am able to pick these specific attributes and link them together through a correlation proof to establish what we can call a "CIP Credential".



In this example, an 'employment credential', consisting of three claimed attributes, has been shared with Acme. Only the claimed attributes of 'salary' and 'name' have been shared with UBS. Such selective disclosure is key to 'data minimization

6.2.2.2 The Relationship-Agent Plane

Because an Identity Owner typically has multiple relationships and thus multiple DIDs, an agent must be able to support multiple DIDs. The Relationship-Agent plane is represented by a "Relationship State Machine," which is implemented with a microledger replicated across the Agents managing a relationship. Agents are classified within the microledger so the other party (or parties) in the relationship know how to regard each agent (identified by a public key). For example, an agent may not be authorized to sign on behalf of a person, but it may be authorized to communicate GPS coordinates or be a message relay. Note that an agent does not have its own DID - DIDs are orthogonal to Agents. Agents can be identified by a local name, or by the public key used in a particular relationship. Each DID-Agent combination requires a separate key. Because an endpoint must be unique per relationship (for the same reasons a private DID is unique per relationship), an agent must also be able to support multiple endpoints. Agents are authorized for certain types of activity for each relationship. This allows the Agent owner to use different agents in different contexts.



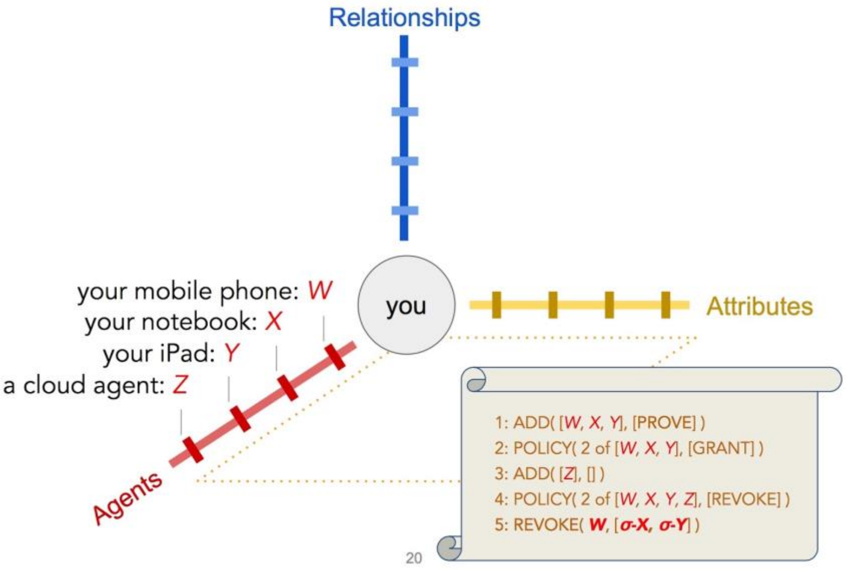
A key unique to the relationship with UBS is established for three of four Agents. These keys can define allowed behavior for each Agent, and also serve to identify the Agent involved in a given transaction. The microledger maintained by the agents on both sides of a relationship keeps track of what has taken place between both parties.

What's a Microledger?

A microledger is a small, ordered, append-only record of events that can be stored in a file or a database table. The microledger contains a log of events plus a merkle tree over those events. The merkle tree makes it easy to detect tampering and allows for safe and efficient propagation of new events. A microledger will also have a state, which is a projection of the events recorded in the microledger. A microledger is replicated across Agents involved in managing a relationship.

6.2.2.3 The Agent-Attribute Plane

The keys Agents have are stored in their own Wallets. Wallets can hold Credentials (verifiable collections of attributes). Agents have special keys that allow them to be able to generate proofs about Credentials they hold. The Agent Authorization Policy is a combination of a ledger smart contract and a cryptographic accumulator that allows for proving in zero knowledge that a device is an authorized device. This allows a person or organization to have a sophisticated recovery policy and to be able to revoke a compromised Agent (as exemplified in the diagram). A person may selectively copy credentials to different agents, effectively limiting which credentials are provable from which agents.


Relationships
you
Attributes
your mobile phone: W
your notebook: X
your iPad: Y
a cloud agent: Z
Agents
1: ADD([W, X, Y], [PROVE])
2: POLICY(2 of [W, X, Y], [GRANT])
3: ADD([Z], [])
4: POLICY(2 of [W, X, Y, Z], [REVOKE])
5: REVOKE(W, [σ-X, σ-Y])
20

The ledger in the diagram above exemplifies the following process flow for revoking a lost phone as an Agent:

1. W, X, Y are established as Agents of the Identity Owner, and are authorized to PROVE (i.e. can generate proofs about the credentials they hold)
2. Policy established that two of the set {W, X, Y} are required to GRANT authorizations Agents
3. Z is established as an Agent of the Identity Owner, with no authorizations.
4. Policy is established that two of the set {W, X, Y, X} are required to REVOKE an Agent
5. Both X and Y are used in a request to REVOKE W, the lost phone

6.2.3 Summary

From this primer, a few observations can already be made. First, despite concerns about the incompatibility of blockchain with GDPR over data storage implications, Sovrin is able to solve this problem by not storing personal data "on-chain" (i.e. on the public Sovrin ledger). Rather, all personal data—including Credentials and Claims—live off the ledger in the Wallets and Agents under an Identity Owner's control. Furthermore, all transactions of this data take place off-ledger in the Agency Layer. Because the only information ever stored on the public Sovrin Ledger does not pertain to "natural persons", such data is is outside the scope of GDPR as the regulation does not apply to entities or things.

Second, so long as an Identity Owner uses a unique DID for each relationship, personal data attributes are only correlated to each other if the Identity Owner authorizes their 'linkage' in a private relationship using a Credential. A Sovrin Identity Owner's ability to 'correlate' attributes on an as-needed and selective basis represents a technical achievement with high significance for privacy regulation compliance.

Moving of data storage and processing off-chain to the extent that Sovrin's model for SSI represents may raise concerns by those familiar with how conventional, public blockchains achieve their immutable and trustless characteristics. It should be noted, however, that Sovrin's protocol utilizes the public Sovrin ledger as an accounting mechanism for maintaining integrity and trustlessness in off-ledger relationships. It should also be noted that Sovrin is able to achieve viable off-ledger transactions without introducing a central authority.

6.3 Fitting Sovrin within Core Data Protection Principles

This subsection evaluates Sovrin's ability to meet data processing principles under Article 5 of the GDPR.

Obligation:

Personal data must be processed lawfully, fairly, and in a transparent manner in relation to the data subject.

Article 6 of the GDPR sets out six lawful bases for the processing of personal data: consent, contract, legal obligation, vital interests, public interest, and legitimate interests.

While none of these lawful bases are new to the GDPR, the "consent" requirement has become more stringent. Consent is defined as "any freely given, specific, informed and unambiguous indication of the data subject's wishes by which he or she, by a statement or by a clear affirmative action, signifies agreement to the processing of personal data relating to him or her." Additionally, the data subject has the right to withdraw this consent at any time and must be notified of this right at the time that consent is obtained. Finally, the data controller must be able to prove that this level of consent was obtained.

In sum, under the high standard set out in the GDPR, valid consent must be:

- freely given
- obtained through an affirmative act of the data subject
- revocable
- provable

Solution:

In the Sovrin model, transactions rely on a technical system of explicit consent for data to be shared in a pairwise relationship. However, it is the onus of the relying party (e.g. the bank) to properly notify the Identity Owner exactly how their personal data (e.g. a Claim) will be used. This does not relate to Sovrin or any technical architecture for that matter, but rather on controller-to-subject communication. It is also important to note that while consent is relied on as a technical means by which an Identity Owner shares data, it may not be the appropriate lawful basis. In the bank example above, for instance, the bank would be using a lawful basis of “legal obligation” or performance of a contract (e.g. opening an account upon request) to collect CIP data.

While lawful basis is more strictly legal and dependent on the context of the relationship, it is important to bear in mind the rights enablement affordances that Sovrin achieves. For example, the Identity Owner can access the information held about them in any pairwise relationship by merely viewing the transparent Claims record. Since DIDs and strict access controls around extra-relationship information sharing make it such that information cannot be collected without literal disclosure from the Identity Owner (i.e. data subject). Regarding data portability, the Identify Owner can switch Agents or Wallet providers at will, by simply moving its claims from one private ledger to another. Objection and deletion can also be achieved through the Identity Owner’s ability to revoke an entity’s access to a particular Claim. Rectification, similarly, would consist of an Identity Owner revoking an out-of-date Claim and re-disclosing the correct Claim to the relying party. (Sovrin’s ability to fulfill data subject rights obligations will be further expounded in Section 6.4).

However, if consent is indeed relied upon by an entity, the Sovrin network architecture allows both parties to easily track the consent along with its revocations and resubmissions. Because both the Identity Owner and the relying party both maintain a log of each digitally signed transaction on their respective copies of their shared private microledger, either can provide proof of consent (or lack thereof) to an auditor. Not only does the individual Identity Owner control the sharing of his personal data, but he is provided with a technical infrastructure that automatically logs consent receipts. This enables an unprecedented degree of transparency to all parties.

Obligation:

(b) Personal data must be collected for specified, explicit and legitimate purposes and not further processed in a way incompatible with those purposes.

The **purpose limitation principle** means that personal data collected for one purpose should not be used or repurposed for a new, incompatible purpose. It is closely related to the lawfulness, fairness, and transparency principle and is designed to minimize correlation risks to the individual data subject and to prevent data controllers and processors from exceeding the limits of lawful processing of personal data. Under the GDPR, a party who is not necessarily a data controller can become one by exceeding the scope of processing that it has been authorized to undertake (at which point it breaks its legal contract with the controller and can incur controller-level liabilities under the Regulation). Thus, it is in the interest of processors to comply with the purpose limitation principle by not exceeding their authorized scope of processing.

Solution:

In the context of Sovrin, a Proof Request can show the purpose(s) for which data is being requested. A Verifier can delete this data once it has executed whatever transaction it was needed for, and then can simply request it again if and when it’s needed. The Verifier can, if necessary, make multiple requests with different purposes, each time making those purposes fully transparent to the Identity Owner who is the subject of the Proof Request. Through the use of a Sovrin Agent, an individual can compare the scope of a Proof Request and its purpose(s) against pre-defined consent parameters in order to more efficiently determine whether the data is being requested for a purpose that is “specific, explicit and legitimate.” In this way, the Agent acts like an intelligent browser, and individuals can implement mechanisms (i.e. Link Contracts) akin to browser settings that help enforce the parameters of personal data disclosure. (This is also forward-looking towards compliance with the forthcoming e-Privacy Regulation).

Verifier: An Identity Owner that wants to rely upon (i.e. obtain access to and be able to evaluate the validity of) the claimed attributes of another Identity Owner

Proof Request: A request on the part of a Verifier for information (i.e. one or more cryptographically established identity attributes) from an Identity Owner

Link Contract: A record of who is sharing data with whom, for what purpose and with what controls on its usage. In Sovrin, a link contract is stored inside a XDI graph. Link contracts are themselves expressed in XDI. Among other stipulations, a link contract can specify how parties prove their identity, how updates are handled, who can use the data and how the data can be used. A link contract is stored on the XDI server (hosted by Agents) at each end of an off-chain relationship, such that the Agents for both parties adhere to the specifications of such a contract.

Obligation:

(c) Personal data must be adequate, relevant and limited to that which is necessary in relation to the purposes for which it is being processed.

The **data minimization principle** has multiple dimensions to it, including limiting the personal data that is collected, processed, and stored. In terms of collection and processing, the same features of Sovrin that give effect to the purpose limitation principle assist in achieving data minimization. Collection is limited by the Identity Owner’s control over the sharing of their data and technical methods that minimize the amount of data that is shared to achieve a given purpose, thereby minimizing the data that is collected and ultimately processed. In Sovrin’s model, the Identity Owner decides precisely which, if any, identity attributes it wants to disclose in the form of the Proofs it produces about the Credentials and Claims in her Wallet.

Solution:

Sovrin is specifically designed to minimize the storage of personal data. Despite concerns about the incompatibility of blockchain with GDPR over data storage implications, Sovrin solves this problem by not storing personal data “on-chain” (i.e. on the public Sovrin Ledger). Rather, all personal data—including Credentials and Claims—live off the ledger in the Wallets and Agents under the Identity Owner’s control. Furthermore, all transactions of this data take place off-ledger in the Agency Layer.

Obligation:

(d) Personal data must be accurate and, where necessary, kept up to date; every reasonable step must be taken to ensure that personal data that are inaccurate, having regard to the purposes for which they are processed, are erased or rectified without delay.

The **accuracy principle** means that data controllers are responsible for taking reasonable steps to ensure that the personal data they hold and process is kept accurate and up to date. The accuracy principle is designed to prevent decisions that have legal or other significant effects on data subjects from being taken on the basis of incomplete or inaccurate information, and is closely related to the lawfulness, fairness and transparency principle. The accuracy principle is also the basis for several of the data subject's substantive rights, including the rights of access, rectification, and erasure.

Solution:

The distributed ledger technology that underpins Sovrin's approach to SSI allows for constant updating and pruning of the public digital record and allows a Verifier to ascertain the status or validity of a Credential or Claim in near-real time by referencing Revocation Registries housed on the Sovrin Ledger.

Revocation Registry

"Revocation Registries" are cryptographic numbers (more specifically, an elliptic curve accumulator) maintained on the Sovrin Ledger by Issuers of revocable digital Credentials that use zero knowledge cryptography to confirm or deny whether a given DID is within or outside of a set of DIDs that have been revoked by a given Issuer

Obligation:

(*e) Personal data must be kept in a form which permits identification of data subjects for no longer than is necessary for the purposes for which the personal data is being processed.*

The **storage limitation principle** is closely related to the purpose limitation and data minimization principles and is designed to prevent unlawful and unauthorized processing and to limit the data security risks posed by the long-term storage and retention of personal data. It is also related to the accuracy principle and the right to be forgotten in so far it is aimed at preventing decisions taken and legal effects for data subjects on the basis of inferences that may be drawn from stale data. Moreover, in the event of a request to be forgotten per the data subject's right to erasure, the acceptable data retention period for a controller or processor may be further shortened from what would ordinarily be a lawful period of time. The exception is for data retained and processed for "historical, statistical or scientific purposes," in which case the public interest may override the data subject's interests.

Solution:

Sovrin enables new data retention strategies that move from a traditional “data management” approach to a “data access” approach where you only use the precise data that you need when you need it, then delete it. You can then repeat if and when necessary. On the whole, this minimizes long-term storage requirements, presents a much lower organizational risk to processors, and facilitates compliance more effectively than the data management approach. Upon receipt of some personal data, the Verifier can execute the transaction they need to carry out and then simply delete the data if it is not needed it anymore. All they need to retain is the DID for the Identity Owner. If the data is needed again, it can be requested again using the secure, private and mutually authenticated Connection with the data subject via their DID.

Finally, because the vast majority of data exchanges in Sovrin happen in the Agency Layer in the context of private encrypted pairwise channels, the storage limitation principle is somewhat self-enforcing. Unique DIDs are created when a pairwise Connection is made and only utilized for purposes of that Connection. When the Connection is terminated (by either party), so is the storage of information in the context of that pairwise relationship.

Obligation:

(f) Personal data must be processed in a manner that ensures appropriate security of the personal data, including protection against unauthorized or unlawful processing and against accidental loss, destruction or damage, using appropriate technical or organizational measures.

The **integrity and confidentiality principle** is fundamentally concerned with data security and the security of processing. It is, in many ways, a threshold principle based on the philosophy that data protection cannot exist without data security. The obligation attaches to all processing, whether by a controller or processor and applies to both external (e.g. hacks) and internal (e.g. employees) security threats. One key technical or organizational measure for data security encouraged by the GDPR is pseudonymization, defined in Article 4 as “the processing of personal data in such a manner that the personal data can no longer be attributed to a specific data subject without the use of additional information, provided that such additional information is kept separately and is subject to technical and organisational measures to ensure that the personal data are not attributed to an identified or identifiable natural person.” In other words, minimize correlation risks to the individual in order to mitigate the impact of a potential security breach.

Solution:

The Sovrin approach promotes the integrity and confidentiality principle by minimizing the frequency and amount of data exchanged in the first place. As demonstrated in reference to the purpose limitation principle, data minimization and storage limitation principles above, Sovrin has made technical and policy decisions to minimize the amount of data that is shared, collected, processed and stored, including through the use of Zero Knowledge Proofs, microledgers, and private

Zero Knowledge Proofs

A zero-knowledge protocol is a method by which one party (the prover) can prove to (i.e. create Zero Knowledge Proofs for) another party (the verifier) that something is true, without revealing any information apart from the fact that this specific statement is true.

DIDs. Moreover, pseudonymization is a pillar of the Sovrin approach. All peer-to-peer communications and Credentials exchanged via Agent-to-Agent messaging are end-to-end encrypted. This means the data is not only encrypted in transit, such as it would be with HTTPS transport-layer encryption, but also the transmitted data is encrypted at rest after receipt by the Verifier, and can only be decrypted by a person or process with access to the Verifier's private key.

Instead of utilizing a public ledger, Sovrin has implemented a public permissioned model, whereby anyone can access the ledger (in terms of reading/writing to it) but only those with permission (e.g. Stewards) can run a node and participate in the consensus protocol to validate the record. This means that there is a level of governance that you could not have with a public permissionless ledger. The Sovrin Foundation implements this governance through a public Trust Framework that includes both security and privacy policies and practices that must be implemented by all Stewards to ensure the security and integrity of the Sovrin public ledger.

Trust Framework

A Trust Framework is a combination of technical "tools" (specifications) and legal/business "rules" (policies) that enables members of a community to trust each other in online transactions. In the Sovrin model, there is a base-level or foundational Trust Framework that is generally applicable to all parties who interact with or utilize the Sovrin Ledger in one form or another (the "Sovrin Trust Framework") as well as higher-level "Domain-Specific Trust Frameworks" that address the unique needs of a specific industry, community, or domain. These Trust Frameworks help address some of the key legal challenges to SSI adoption, including in relation to the GDPR.

Obligation:

(2) The controller shall be responsible for and be able to demonstrate compliance with the GDPR principles.

The **accountability principle** is closely related to the lawfulness, fairness and transparency principle and requires, at its core, an effective means to demonstrate compliance. Data controllers are primarily responsible for compliance with data protection obligations under the GDPR (see Section 2.1.2). Therefore, in order to determine the degree to which the accountability principle is given effect, we have to evaluate compliance and the auditability of that compliance in light of who the controller is in any given transaction.

Solution:

The Sovrin model for SSI promotes accountability through its multi-pronged approach to governance. At a technical level, the following elements provide an unprecedented level of accountability for controllers:

- The formation of private pairwise channels of communication
- The mutual authentication that happens through Connections

- An automatic, digitally signed, auditable record of each transaction taking place on each microledger

The Sovrin public distributed ledger allows for a fully transparent and auditable record of the public events that transpire in the Sovrin ecosystem. More importantly, in Sovrin, Trust Frameworks include legal and commercial policies that reinforce the accountability achieved by technical design decisions and introduce additional mechanisms for governance and accountability. Finally, because accountability attaches to the controller (usually the Verifier in a Sovrin-style exchange of Credentials), there are strong incentives for corporations and organizations acting as Verifiers to adopt the Sovrin solution as a means to demonstrate compliance.

Obligation: Privacy by Design and Default (Article 25)

In addition to the seven core principles outlined by Article 5, the GDPR imposes a general obligation to implement technical and organizational measures that give effect to all of these principles of data protection in an integrated fashion, an approach known as privacy by design and default. Rather than the conventional approach that companies have taken—where privacy and data protection were treated as afterthoughts left to compliance departments—privacy by design and default asks organizations to consider privacy and data protection from the start and build them in as key features of their products and services (this requires a degree of legal engineering of product offerings).

Solution:

Sovrin achieves privacy by design by establishing the following defaults within its protocol:

- *Pseudonymity by default* – the use of DIDs and public keys to achieve trustless, pseudonymous identity.
- *Private Agents by default* – the private data of Sovrin users, encrypted or otherwise, is not stored on Sovrin's ledger.
- *Selective disclosure by default* - cryptographic zero-knowledge proofs minimize the number of identity attributes that need to be shared to adequately establish identity for a given context, and only those that have been provided with the proper cryptographic keys may access these identity attributes.
- Robust governance mechanisms - including the use of a permissioned ledger and various Trust Frameworks.

6.4 Sovrin's Compliance with Data Subject Rights

Article 17 - Right to Erasure

Regulatory Obligation:

Data subjects have the right to have personal data erased under certain circumstances, including where the personal data is no longer necessary for the purpose(s) for which it was originally collected; where the personal data was processed on the basis of consent and the individual withdraws their consent; where the individual objects to continued processing of personal data processed on the basis of legitimate interests and there is no overriding legitimate interest to continue to process it; where the data was processed unlawfully; where required by law; or where the data was processed to offer information society services to a child. This right is also designed to give effect to the accuracy principle and the lawfulness, fairness, and transparency principle.

Sovrin's Solution:

Any PII or other data that an Identity Owner may choose to associate with a DID is either (a) stored some someplace off-ledger (e.g., by a Sovrin app or agent) or (b) shared peer-to-peer with relying parties under a link contract that enables the Identity Owner to assert the right that private data be forgotten. With Sovrin's pairwise Connections and microledger technology, each individual Identity Owner will have a separate record of every data sharing event for every individual or entity they ever interacted with. An individual Identity Owner can use the records generated by their microledgers and pairwise Connections to prove what information they shared, when they shared it, and with whom they shared it to make a request to have certain data erased. Each Identity Owner in a pairwise Connection can use their respective Agent to seamlessly request the erasure of their data by the peer on the other end of the Connection. Such a request itself becomes a provable event via their shared microledger.

With this design, Sovrin infrastructure is ideal for implementing the right to be forgotten because *a* signed link contract stored with a Sovrin Agent gives an Identity Owner his/her own tool for exercising the right to erasure at any time. The Identity Owner can simply instruct his/her Sovrin agent to send a digitally signed "forget me" request to the relying party. That request includes the link contract under which the private data was shared (which also includes the Identity Owner's DID). Once the relying party verifies the digital signature on the 'forget me' request, the relying party will have everything it needs to delete the private data and return a signed acknowledgement (by either Sovrin agent or link contract) that it did so back to the Identity Owner, plus the audit trail necessary to prove compliance to a regulatory body. Because the request for erasure itself can be a provable event, if the erasure is not made the individual Identity Owner can use this proof of sharing and requested erasure to take action against a peer who continues to hold and process their data unlawfully

Articles 13 & 14 - Right of information

Regulatory Obligation:

Data subjects have the right to be informed about the collection and use of their personal data. This right is designed to give effect to the lawfulness, fairness, and transparency principle. This means that at the time when personal data is collected from an individual, the individual must be provided with information regarding the identity of the party who is collecting or requesting the data, the purposes for collecting or processing the data, how long the data will be retained, and who the data will be shared with (if anyone). This information must be provided in a concise, transparent, intelligible and easily accessible format, using "clear and plain language." These requirements do not apply where the data subject already has the information.

Sovrin's Solution:

When a Proof Request is received by an individual Identity Owner, it contains the details describing the requesting party (i.e. the Verifier) and the data that is being requested. The Identity Owner keeps their own record of this request through their Agent. If a given Verifier has already identified and authenticated itself to the Identity Owner in the process of establishing a pairwise Connection and making a Proof Request, the individual can be said to already have the information, meaning these obligations are automatically fulfilled for a repeat of the same Proof Request. In Sovrin's implementation of SSI, the Identity Owner has full control over their identity attributes. This means that the Identity Owner, as the data subject, will at all times have this information, which means the right of information is automatically realized in Sovrin's implementation of Self-sovereign Identity.

Article 15 - Right of Access

Regulatory Obligation:

Data subjects have the right to request and obtain access to their personal data and certain other supplementary information. This right is also designed to give effect to the lawfulness, fairness, and transparency principle. The right of access gives the individual data subject a right to obtain confirmation that their data is being processed, access to their personal data, and other supplementary information that would typically be provided to the individual via a privacy policy per their right to be informed. Upon the request of the data subject, this data must be provided without delay and at the latest within one month of receipt of the request. As a best practice, "Where possible, the controller should be able to provide remote access to a secure system which would provide the data subject with direct access to his or her personal data." Finally, a party who receives a request for access has a duty to use "reasonable means" to verify the identity of the person making the request.

Sovrin's Solution:

With Sovrin's pairwise Connections and microledger technology, each individual Identity Owner will have a record of every data sharing event with every individual or entity they ever interact with. They will therefore be able to prove precisely what they shared, with whom, and when. Not only can the Identity Owner use this information to demand access, but the recipient of a subject access request can use their private Connection (and their private key dedicated to this Connection) to authenticate the identity of a peer making this request. This solves the problems of both an Identity Owner (i.e. data subject) needing access and a Verifier (i.e. data controller) needing to provide it. This private channel can also be used by the Identity Owner to request access to any other personal data generated by the Verifier, or personal data disclosed to or shared with others.

Article 16 - Right to rectification

Regulatory Obligation:

Data subjects have the right to have inaccurate personal data rectified or incomplete data completed, including by way of supplementary information. This right is designed to give effect to the accuracy principle and to prevent decisions with legal or other significant effects on the data subject from being made on the basis of inaccurate or incomplete information.

Sovrin's Solution:

As described above, an individual Identity Owner in the Sovrin ecosystem can use the microledgers generated in the context of each pairwise Connection to prove exactly what information was shared, when it was shared, and with whom it was shared. Moreover, through its respective Agent, either party in a pairwise relationship has a seamless means to request a correction or other modification to their personal data from their peer on the other end of the Connection. Such a request itself becomes a provable event via their shared microledger, such that if the changes are not implemented the requesting party (who is the subject of a rectification request) can use this proof of sharing and requested rectification to enforce its rights against the peer (acting as the data controller in respect of that data). This mechanism also helps reinforce the accountability principle.

Article 18 - Right to Restrict Processing

Regulatory Obligation:

Data subjects have the right to restrict processing by a data controller where the accuracy of the personal data is contested (while accuracy is under review); the processing is unlawful but the data subject requests restriction rather than its erasure; the controller no longer needs the data for processing but the data is required for the establishment, exercise or defense of legal claims; or where the data subject has objected to processing and the controller is considering whether its legitimate grounds override those of the individual data subject.

Sovrin's Solution:

As covered extensively with respect to the rights to rectification and erasure, the Sovrin solution—primarily through the use of private pairwise Connections and corresponding shared microledgers—allows the Identity Owner (i.e. data subject) to easily and verifiably exert this right over any personal data shared with a Verifier in the context of that pairwise channel. Moreover, in respect of the public ledger, to the extent that no personal data is ever stored on the public ledger, there is no obligation to restrict processing in respect of the Sovrin Ledger.

Article 20 - Right to Data Portability

Regulatory Obligation:

Data subjects have the right to view, access and obtain their personal data from one data controller and to reuse or transmit that personal data to another controller for their own purposes. According to the Article 29 Working Party, the right "represents an opportunity to 'rebalance' the relationship

between data subjects and data controllers, through the affirmation of individuals' personal rights and personal data concerning them."
The right applies where three conditions are met:

1. the personal data was provided directly by the data subject to a data controller;
2. the processing was based on the consent of the data subject or the performance of a contract;
3. where processing is carried out by automated means.

The data controller must provide the personal data in a "structured, commonly used and machine-readable format." Where requested by the individual and technically feasible, the data controller may be required to transmit the data directly to another controller. Finally, data portability and the transfer of personal data thereunder requires that the transferring party take reasonable measures to verify the identity of the receiving party.

At first glance, data portability feels as though it has little to do with the protection of the personal data of a natural person, perhaps apart from a tangential connection to the transparency principle. Rather, this right encapsulates the GDPR's second objective—namely, to promote commerce and growth by enabling the free movement of data across the EU. Data portability is mutually beneficial for the individual who can more conveniently access and move her data across service providers or organizations with ease, as well as the commercial entity who can more readily absorb and onboard new users and customers. Data portability also fosters competition when individuals are not locked into the products or services of one service provider and can switch providers without many impediments. This is in line with other pro-competition initiatives in the EU such as the Open Banking initiative.

Sovrin's Solution:

Sovrin's version of Self Sovereign Identity is the ultimate data portability tool. Just as in the physical world, I can take my documents with me from one location to another and present them to verifying or requesting parties as needed, Sovrin enables the same functionality in the digital realm. Because the Identity Owner holds Credentials and other personal data in a portable digital Wallet, the Identity Owner is free to use and move these credentials as desired. This in squarely in line with the Article 29 Working Party's view that primary aim of data portability is enhancing individuals' control over their personal data and making sure they play an active part in the data ecosystem. Finally, Sovrin offers a strong technical solution in respect of a transferring party's ability to authenticate the party receiving the data that is subject to a transfer request per the right data portability.

Article 21 - Right to Object

Regulatory Obligation:

Data subjects have the right to object to three kinds of processing:

1. direct marketing;

2. processing based on legitimate interests of the data controller or the performance of a task in the public interest or the exercise of official authority;
3. processing for research or statistical purposes.

The right to object to direct marketing is absolute in the sense that the individual need not demonstrate any grounds for this objection and as soon as the individual objects, the processing must cease. In the case of the other two bases for objecting, the controller must cease processing unless it can demonstrate compelling legitimate grounds which override the interests of the data subject or the processing is for the establishment, exercise or defense of legal claims.

Sovrin's Solution:

As with the right to restrict processing, Sovrin Connections and microledgers enable Identity Owners to easily and verifiably exert this right over any personal data shared with a Verifier.

Article 22 - Rights in respect of automated decision-making and profiling

Regulatory Obligation:

Data subjects have the right "not to be subject to a decision based solely on automated processing, including profiling, which produces legal effects concerning him or her or similarly significantly affects him or her."

The only exceptions are:

1. where this kind of automated processing or profiling is necessary for entering into or the performance of a contract between the data subject and data controller;
2. the decision-making is authorized by applicable laws that include suitable measures to safeguard the rights of the data subject;
3. the decision-making is based on the data subject's explicit consent.

Sovrin's Solution:

Sovrin Connections enable the data subject to easily and verifiably exert this right via Link Contracts. While it might appear that Sovrin Agents would themselves be subject to this automated decision-making rule, in fact *Sovrin Agents only act on behalf of the data subject and always remain under the control of the data subject*. Thus a Sovrin Agent does not operate under the control of an external data controller that would be subject to this rule.

6.5 Alternatives to Achieving Secure, Trustless, Off-ledger Interactions

As our case study of Sovrin describes, using a distributed ledger as an accounting and transparency mechanism while moving the storage and processing of personal data off-chain can be an effective approach for achieving compliance with data privacy regulations, while retaining the trustless and decentralized characteristics of 'vanilla' blockchains. Sovrin's SSI model is not the only means for achieving this type of solution, however. Accordingly, in this section we provide a cursory summary of alternative approaches off-ledger transactions, intend to ‘round out’ our treatment of off-ledger-type solutions for data privacy regulations. This summary includes introductions to (i) the more general concept of state channels, of which Sovrin’s ‘Agency Layer’ is a particular implementation and (ii) the technique of Trusted Execution Environments and (iii) the use novel use of TEEs by Teechain.

6.5.1 State Channels

State channels are private, two-way pathways opened between two users that want to communicate with each other in the form of transactions, off chain. These communication pathways allow for the same data integrity and trustless properties of on chain transaction processing for transactions that take place off chain by utilizing a ledger or set of ledgers as an accounting and validation mechanism for off chain transactions. Payment channels are a subset of state channels where the state being tracked is payment.

Both participants in a state channel signs off chain transactions with their private key (or equivalent for other cryptographic schemas) to establish that they are authorized and undeniably true. Because of their off-chain nature, such transactions can take place without the potential for latency introduced by on chain consensus mechanisms. State channels usually have a limited lifespan, predetermined based on time or amount of transactions carried out. Participants can also close their channel by providing the last-updated state of transactions within the channel to the blockchain. If the other participant agrees that the state provided is indeed the final one, they let the channel close. If not, they can upload their version of the most recent state and since all transactions are signed and time stamped, such a dispute will always be won by the honest party.

The contents of transactions that take place within a state channel are private and known only to its participants - this is highly significant - by utilizing state channels, sensitive data is only shared with those that need to have access to it. Once a channel closes, the transaction history of the channel is recorded on an immutable ledger as a hash, which can be referenced by those who have access to said transaction history (i.e. the parties involved in a state channel) to establish the validity of a given record of this same transaction history. In addition to significantly reducing the amount of information that individual nodes need to store, state channels also reduce the number of transactions that need to be processed by nodes, as the set of transactions that took place in a state channel can be processed as a single transaction when the state channel closes.

For situations in which the off-chain transaction state of more the two parties must be monitored, multi-hop state channels have been proposed. Such channels processes transactions in a manner that involves multiple participating nodes as opposed to only the two nodes sending and receiving a payment or evaluating a state. Successful multi-hop payment channels were demonstrated by the Bitfury Group in May of 2017 on the Bitcoin Lightning Network. Multi-hop state channels, which are more complex, are still in the developmental phase. Until recently, proposals for state channels in general had seen little adoption due to their incompatibility with current blockchain protocols, practical limitations, and implementation complexity. New conceptions of how they can be achieved, however, by efforts such as Teechain, Sovrin, and IBM’s deployment of Hyperledger Fabric are moving this concept from the theoretical to the feasible.

6.5.2 TEEs

A Trusted Execution Environment (TEE) is a trusted region on a processor isolated from, but running in parallel to, the main operating system of a device that utilizes both hardware and software to protect the integrity and privacy of data. TEEs are intended to guarantee robustness in the presence of an attacker who has full control of the hardware and has compromised privileged software, including the OS and hypervisor of a virtual memory.

By only storing, managing, and using encryption keys within a TEE, data, communications, and cryptographic operations can be secured with no opportunity for eavesdropping. This is particularly relevant for increasing the security of state channels and other privacy-by-design blockchain techniques (such as Audit-enabled Privacy, discussed in Section X.x) because cryptographically-established access permissions are only as secure as the security of the private keys that endow access. The most advanced TEE implementations, such as Intel's Software Gaurd Extensions (SGX), embed devices with unique identities via roots of trust - a set of functions, part of a trusted computing environment, that are always trusted by a processor's OS. These enable key stakeholders (i.e. network operators, data processors) to identify whether the device they're interacting with is authentic - otherwise known as remote attestation functionality.

It is worth noting that Sovrin's Agency Layer allows for similar environments without the use of special hardware with OS-level software to establish locations for secure data processing and storage. This allows for Sovrin Agents to be run from personal computing devices, such that the processing and storage of data in the Sovrin network can occur in a more decentralized manner - i.e. not consolidated across the purchasers of expensive TEEs.

6.5.3 Teechain

Teechain is an innovative off-chain payment protocol that demonstrates a novel approach for (i) utilizing TEEs to address shortcomings of previous payment channel deployments and (ii) further increasing the resiliency of TEEs. Of particular significance is Teechain's use of TEEs to enable asynchronous blockchain access for payment channels, which allows for validated offline transactions and reduces the latency for payment channel transactions.

Most payment channel deployments require synchronous access to the blockchain. This means that the parties in a channels must leave the channel at the same time, with a hashed record of the state of the channel being recorded as a transaction on the blockchain. Although a user on such a deployment can, at any time, settle the channel for all parties, it is possible for a user to request the channel to be settled at a depreciated state, using the 'previous' capabilities that accompany said depreciated state (i.e. stale state exploit). To prevent such attacks, many solutions require users to monitor the blockchain and react to such misbehavior by providing an authenticated timestamp of the 'non-stale' state of the channel that was recently closed in such a manner, which places a burden on users and allows for invalid records to be written on-ledger, even if temporarily. Addressing such erroneous entries places a burden on the blockchain network, which decreases its performance.

Teechain avoids necessity of such monitoring by use of collateral in the form of on-blockchain deposits to secure payment commitments within a given channel. The collateral is maintained by TEEs, allowing users to dynamically move funds between payment channels - important for enabling multi-hop channels. Because the TEEs protect the internal channel state and release it only upon channel termination, they ensure that users cannot launch attacks by using a stale state exploit. This approach is intend to avoid such attacks on payment channels, and simplifies the protocol such that network performance is improved.

Teechain mitigates known failures of TEEs with two strategies: (i) backups to persistent storage and (ii) a novel variant of chain-replication, both of which are intended to increase the resiliency of data contained in TEEs by establishing additional sources of data validation. Although Teechain's protocol is intended to be able to function on any TEE, their first implementation uses Intel's SGX. Although Teechain's solutions were developed for payment channels and not state channels, which are more complex, their techniques demonstrate how TEEs can be used to secure off chain transactions, and enable offline off chain transactions. It should be noted, however, that Teechain's payment channel solution, offline-enabled and efficient as it is, only allows only for payment-type transactions - not the more complex types of off-ledger interactions enabled by Sovrin-type state channel.

6.6 Sovrin Alternatives for KYC-AML Compliance

This section is comprised of two parts. 6.6.1 introduces a conflict between Sovrin's protocol and KYC-AML and then describes a modification Sovrin's protocol that could resolve this conflict. 6.6.2 introduces an enterprise-oriented distributed ledger model - Hyperledger Fabric - which was designed for KYC-AML compliance in mind, while Sovrin was not. Included in this introduction is an overview of two features of Hyperledger Fabric that are uniquely suited for achieving KYC-AML compliance.

6.6.1 Potential Conflict between Sovrin and KYC-AML

As discussed in Section 6, when one of the Sovrin Identity Owners in a relationship requests the relationship to be closed, access to data shared over the course of this relationship disappears. This aspect of the Sovrin protocol represents a potential conflict with record-keeping requirements specified by CCD, specifically the requirement of hold records of personal information for five years after an account is closed. We propose on approach identical to Sovrin but with one key difference as a solution for this conflict: In relationships where AML / KYC compliance is required by one or both Identity Owners, a second, *compliance* microledger is also maintained across the Agents used by both Identity Owners for said relationship. This compliance microledgeer would maintain a record of the personal information attributes specified by CCD under its record-keeping requirement - name, date of birth, address, and ID number. Unlike the ‘primary' microledger used in conventional Sovrin relationships, this compliance ledger would persist for five years after the termination of the relationship.

6.6.2 Hyperledger Fabric

Hyperledger Fabric is a business blockchain framework hosted by the Linux Foundation. Fabric is actually related to Sovrin, as the latter is a protocol running off of Hyperledger Indy - an identity blockchain framework and the brainchild of Sovrin foundation. Fabric was developed to address the unique set of considerations (KYC-AML regulations, specifically) faced by enterprise applications. In addition to utilizing state channels, Fabric proposes two techniques of practical significance for achieving KYC-AML compliance within a distributed ledger network.

The Zero Knowledge Asset Transfer feature of Hyperledger Fabric allows transactors to issue assets and request transfer of their assets without revealing anything to the public ledger for the assets being exchanged. Unlike other privacy-preserving asset management systems for blockchain, ZKAT is tailored to the needs of enterprise networks. In particular, the auditability of the privacy-preserving transactions within Fabric is a crucial differentiator from other schemes available. This feature is known as *Audit-enabled privacy* - each user is assigned a specific auditor who is entitled to unlimited access to all the transactions of that user. The auditors are passive; they may come in afterwards and extract the confidential information of all transactions the audited user is involved in, but they cannot access the data for any other party. Because Audit-enabled privacy allows for KYC and AML compliance within blockchain networks by granting auditors, and auditors only, access to personal data for specific users, data processors (besides auditors) and network operators are not subject to the data privacy regulations that come with handling such data.

With Fabric, private transactions (i.e. state channels) can be combined with anonymous client authentication (i.e. Audit-enabled privacy) to avoid leaking the connection between the identity of the transaction's creator and the ledger stored (hashed) data.

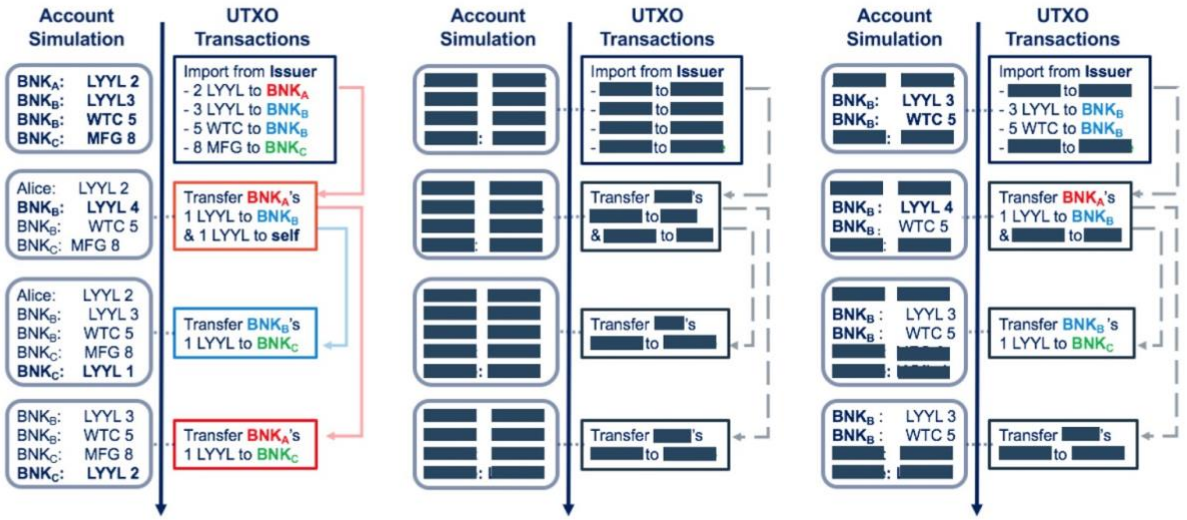


Figure 5: Audit-enabled privacy

In the figure above, the access to on-ledger financial data provided by distributed ledger protocols to an auditor is shown across three different privacy mechanisms. The model on the left has no privacy mechanism. The model in the middle supports full privacy, concealing the identities of transaction participants and traded assets details. The model on the right demonstrates the secure auditing capability of ZKAT - the auditor assigned to a specific Bank is allowed unrestricted access to all transactions involving this bank.

Annex 1: GDPR Terms and Additional Requirements

Personal Data vs. Non-Personal Data

Personal data in the GDPR is about the ability to attribute a piece of information to a person. This means that whether a particular type of information is classified as "personal data" depends on the context of the personal data in the dataset and an entity's ability to use the data to link to a person.

> **Personal data** is "any information relating to an identified or identifiable natural person ('data subject'); an identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier…"
>
> - *Article 4(1) GDPR*

The GDPR's definition of "personal data" gives some examples of data that commonly reaches the threshold:

- Name and surname
- Home address
- Identification number (e.g. social security, driver's license)
- Location data
- Internet Protocol (IP) address
- Cookie ID
- Online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person

Here are a few other common contenders for personal data to this list:

- Bank or credit card details
- Transaction history
- Health or biometric information
- CCTV footage

These data are almost always classified as "personal data" because they can easily identify a legal person with little to no extra information. Other data, such as transaction history, personal preferences, and other economic/social/cultural factors can be used in combination with other associated data to easily narrow down and identify a specific person within or outside a dataset.

Why is an IP Address personal data?

IP address is a number assigned to any device connected to the Internet. IP addresses are public and accessible to anyone and any server you communicate with. An IP address can determine a device's geolocation (city, state, and sometimes zip code) and its Internet Service Provider (ISP). While IP address can be spoofed (through Tor or a VPN) and cannot actually identify the person behind the device, it vastly narrows down the list of possible people and can be used in combination with other information to identify you.

Before moving on, it is important to understand the last part of the GDPR's definition about "identifiable" data being "directly or indirectly" related to a person. Here the GDPR is stressing that you may not even have the ability to identify the person through your data, but the ability to identify would be easy given another piece of data.

Example

Say you have a data set where each row corresponds to a transaction (e.g. sweater_10€), and each transaction is linked to a customer's account number. In this situation, the account number is directly "personal data" because it references an individual person; the transaction would also be "personal data" because it could be used to identify the account number and therefore the person behind the account. That is what the GDPR calls "indirect" identification. However, if you give an analytics company just the list of transactions without the account numbers, then the transactions no longer become personal data.

Special Categories of Personal Data

"Special category" data are a subset of personal data, defined under Article 9(1) as data that reveals:

- racial or ethnic origin
- political opinions
- religious or philosophical beliefs
- trade union membership
- genetic data
- biometric data for the purpose of uniquely identifying a natural person
- data concerning health
- data concerning a natural person's sex life or sexual orientation

Note that the data does not necessarily need to reference a special category to be labeled a special category. Similar to the idea of indirect identifiers, if a piece of data can be used to infer a special category, then it should be treated with special category protection.

Business Requirements for Processing Special Category Data

If you process special categories of personal data, then you need to keep in mind the following requirements:

1. You're not allowed to process special categories at all unless one of the following apply:[70]
    a. You have explicit consent from the person
    b. You are under a legal obligation
    c. The data is public information
    d. Processing is in either the public interest or the person's vital interests

2. You need to keep a record of your processing activities, even if you have fewer than 250 employees[71]

3. You cannot use special category data in any automated decision-making processes such as profiling (i.e. no human input), unless:[72]
    a. You have explicit consent, or
    b. The processing is necessary for reasons of substantial public interest

4. If you process special category data on a large scale, then:
    a. You need to have a Data Protection Officer (see our DPO guidance document for more details).
    b. You need to conduct Data Protection Impact Assessments for any technologies involved in this processing activity

EU Representative

Do I need an EU Representative?

Regardless of whether you're a processor or a controller, you need an EU Representative if you don't have an office/establishment in the EU and you do one of the following:

a) Regularly monitor the behaviour of people in the EU (for example, tracking them for analytics, web marketing, or online advertising purposes).
b) Collect/process/store a large amount of data that relates to people's racial or ethnic origin, sex or sexual orientation, biometric or genetic information, health data, or political/religious/philosophical beliefs.

[70] Article 9(2)
[71] Article 30(5)
[72] Article 22(4)

What does an EU Representative Do?

Role: An EU Representative acts as your company's point of contact for European regulators and data subjects regarding data processing and GDPR compliance issues. The Representative needs to be able to respond to any data subject requests and Supervisory Authority mandates, so they should at least have a rudimentary knowledge of data protection requirements and appropriate response protocol.

Responsibility: while the EU Representative is your main liaison in the EU, it's important to note that you (the controller or processor) still bear full liability over any data breach or misuse of data by your company.

How do I go about appointing one?

You can appoint an EU Representative from within your company or hire an external party to represent you. Either way, this representative should be physically located in the EU. Your EU Representative should be "explicitly designated through written mandate". While further explanation has not been given, we suggest you write down the contact information of your EU Representative on your company website's Privacy Policy.

Data Protection Officer

Do I need a DPO?

Article 37 of the GDPR says you need a DPO if:

1. You're a **public authority**.
2. Your **core activities** involve **regular or systematic monitoring** on a **large scale.**
3. Your core activities consist of processing **special category data** on a **large scale**.

For most people, these three conditions shed little light on whether they need a DPO because all the terms in **bold** are highly ambiguous. So let's rephrase the above in a more understandable way. You need a DPO if at least one of the following apply on a large scale:

1. Web analytics, targeted advertising, or online marketing is a main part of your business.
2. You're a local government body, a trust, university, publicly-funded organisation, or similar type of organisation.[73]
3. At least some of your company's key offerings and activities to other clients involve processing personal data to track or profile people (e.g. for behavioural advertising), and you do this on a regular or recurring basis.

[73] The full list of types of public bodies defined by the UK government can be found here: https://www.gov.uk/government/publications/public-bodies-2016

4. You process personal data that includes or infers information about someone's:[74]
    a. racial or ethnic origin
    b. political opinions
    c. religious or philosophical beliefs
    d. trade union membership
    e. genetic or biometric information
    f. sexual characteristics or sexual orientation
    g. criminal history

Finally, to understand the "large scale" condition that affects (2) and (3), the Article 29 Working Party suggests that you consider the volume of personal data you process, the number of individuals whose personal data you process, and how often you process these types of data. Unfortunately, there's no bright line or quantitative data about whether your company meets the "large scale" threshold. Here are some examples of instances we'd advise you to hire a DPO:

a) Your company has multiple employees that process special category data or perform online tracking/profiling as a full-time job for your company.
b) Your company has multiple locations and use loyalty cards or monitor the internet usage of customers using your wifi (or if you sell that information to another company).
c) Your company has multiple locations and tracks the geo-locational data of its customers at each location (e.g. by requiring customers to "check-in" before accessing your wifi).

The DPO's job

Tasks:

- Oversee the company's data protection compliance programme and assist (or lead) the overarching data management strategy.

- Act as the lead and/or advisor when conducting implementations, such as Data Protection Impact Assessments (DPIAs), records of data processing, rights enablement, and data protection by design and default.

- Managing data breach mitigation and response protocol and informing the Supervisory Authority (along with EU Representative if relevant).

- Ensuring that the company's record of processing activities is up-to-date, and that the contents of the record comply with the GDPR (both in terms of content and lawfulness of processing).

- Ensure that all employees that interact with personal data undergo GDPR education compliance training and are familiar with the company's policies, such as the Data Protection Policy and Data Breach Protocol.

[74] This list comes from the GDPR's definition of "special categories" of data, which can be found in Article 9(1) of the text. https://gdpr-info.eu/art-9-gdpr/

- Inform and advise the executive team on GDPR and data protection compliance.

National Derogations

The threshold for requiring a DPO can change depending on the Member State you're operating in. As of January 2018, we're only aware of one country (Germany) to establish a lower threshold for requiring a DPO.[75]

In **Germany**, you need a DPO if:[76]

- At least twenty employees process personal data as part of their regular activities (e.g. HR)
- At least ten employees regularly conduct automated processing of personal data
- Your business conducts any high-risk processing that trigger the need for a Data Protection Impact Assessment (DPIA)
- If you process personal data commercially for the purpose of either marketing or opinion research.

Personal Data Transfers Outside the EEA

Overview

The GDPR imposes strict requirements on companies processing or transferring personal data outside the European Economic Area (EEA) or countries with "adequacy". This section helps you:

- Understand which countries you can transfer personal data to without legal safeguards.
- Evaluate the three legal options for transferring personal data to a "third-country".
- Determine the right legal approach for your business and implement it effectively.

Introduction

Many – if not most – companies end up processing personal data outside the country in which the data is collected. This is because the data processing lifecycle rarely stays fully in-house; companies instead choose to outsource some or all of their processing, storage, or another component to a different company.

But when you transfer personal data to a company like Amazon Web Services, Salesforce, Microsoft, or a smaller vendor, you need to make sure that you are doing it in a GDPR-compliant manner. The GDPR allows you to transfer personal data of EU residents outside the EEA, but it imposes strict requirements on how you do so, and the legal safeguards that must be in place before making the transfer.

[75] IAPP has a list of other country derogations for the DPO requirement. However, at this point in time, it may be too early to tell decisively what the end result will be. Link: https://iapp.org/news/a/gdpr-conundrums-the-data-protection-officer-requirement/.

[76] See: https://www.twobirds.com/en/news/articles/2017/germany/germany-is-the-first-eu-member-state-to-enact-new-data-protection-act-to-align-with-the-gdpr;

Countries that Do Not Require Additional Safeguards

Here's a list of countries that are either in the EEA or have obtained "adequacy" from the European Commission. It is perfectly fine to process personal data in these countries or transfer personal data to an entity in any of these countries without needing further legal safeguards.

<u>EEA Countries</u>

*(countries in EEA but not in EU in **bold**)*

- Austria
- Belgium
- Bulgaria
- Czech Republic
- Cyprus
- Denmark
- Estonia
- Finland
- France
- Germany
- Greece
- Hungary
- **Iceland**
- Ireland
- Italy
- Latvia
- **Liechtenstein**
- Lithuania
- Luxembourg
- Malta
- Netherlands
- **Norway**
- Poland
- Portugal
- Romania
- Slovakia
- Slovenia
- Spain
- Sweden
- United Kingdom

Countries with "Adequacy" Designation

- Andorra
- Argentina
- Canada (commercial organisations)
- Faroe Islands
- Guernsey
- Israel
- Isle of Man
- Jersey
- New Zealand
- Switzerland
- Uruguay
- United States of America (limited to companies registered to Privacy Shield)[77]

Transferring Personal Data to a "Third Country"

If you are transferring personal data to a country not on either of the above lists, then you will need to use one of these specific legal mechanisms to ensure compliance with Article 46 of the GDPR.

Binding Corporate Rules

What are they?

Binding Corporate Rules (BCRs) are legally binding internal codes of conduct for organisations within a corporate group or organisations participating in a "joint economic activity", such as a joint venture or engagement between multiple businesses. They regulate the processing of personal data within and between companies in a company group, including across borders.

Should I use them?

BCRs are useful for international organisations with multiple subsidiaries or affiliate groups under a single umbrella. This way, the organisation only needs to make one BCR for each subsidiary to be covered, as long as they participate as a signatory to the BCR. It is also important to note that both controllers and processors can use BCRs. While BCRs have greater up-front costs in both time, money, and transparency to authorities, they tend to pay off in the long run. This is because they tend to last longer, they cover other necessary policies (e.g. data protection policy), and they often give regulators more trust in your company's compliance regime than other legal safeguards such as standard contract clauses.

What needs to be in them?

GDPR Article 47 outlines specific content requirements for BCRs, including:

[77] You can find a list of companies registered to Privacy Shield here: https://www.privacyshield.gov/list.

- The list of parties covered under the BCR as well as their roles in the data processing
- The material scope of the BCR, i.e. the specific types of personal data processed, the categories of the recipients, and the countries to which personal data will be transferred.
- An explanation of how data protection principles will be applied under the BCR (e.g. data protection by design, security, data minimization, legal bases for processing, etc.)
- Data subject rights, how they can exercise those rights, and how to lodge a complaint
- Complaint procedures
- The distribution of liability among each party
- Mechanisms for reporting and recording changes to the BCR
- Mechanisms for demonstrating compliance by each Member to the Supervisory Authority

More information about BCRs from the Article 29 Working Party can be found here: https://ec.europa.eu/newsroom/just/document.cfm?doc_id=48799

How do I implement a BCR?

You need to submit your BCR through an application form to the Supervisory Authority in the EU member state in which you are established (or the primary member state in which you do business). Only one BCR application form need be submitted by a group, so you should choose which entity will serve as the lead and liaison between the group and the Supervisory Authority. The application form can be found and downloaded in the Supplementary Guidance section of the Hub.

Standard Contractual Clauses

What are they?

Standard Contract Clauses (SCCs, aka Model Contract Clauses) are data protection clauses written and approved by the European Commission and incorporated into your Data Processing Contract. They can be either Controller-Controller clauses (whereby the “data exporter” and “data importer” are both controllers and can make decisions on how data is processed) or Controller-Processor clauses.

Should I use them?

SCCs are quicker to implement than BCRs because, unlike BCRs, you don’t need to wait for review and approval from the Supervisory Authority. You might also want to use SCC if you only want clauses for international transfers; BCRs require you to fill out both broader and more detailed information on your overarching data processing activities. However, because of the added transparency and accountability required in a BCR, EU regulators tend to hold greater trust in BCRs as demonstrators of an organisations overall compliance regime and commitment. Finally, an SCC might make more sense for a company that doesn’t plan on making many international transfers of personal data and only wants clauses for limited and specific client engagements.

What needs to be in them?

The European Commission provides the exact text for the SCC. This includes the names of the contracting parties, obligations of each party, mediation, and jurisdiction, and space for you to describe information on the types of data processed and their recipients.

Note: these SCC templates were created under the Data Protection Directive, before the GDPR came into effect. While the European Commission still has the original templates published on its site, we can expect that they will be updated in the near future. We will do our best to monitor this and let you know if any substantial changes occur, but you should also be aware that the current SCC templates may not last long.

How do I implement an SCC?

If you are relying on any of the European Commission sets of SCCs, you can't change the clauses in any way (except to add another party). However, note that there are specific sections in the Appendices/Annexes where you will need to fill out information. You can either use SCCs as stand-alone-contracts or as a part of a larger data processing service agreement. The European Commission's SCC example clauses for both controllers and processors can be found in the Supplementary Guidance section of the Hub.

Approved Codes of Conduct

What are they?

Approved Codes of Conduct are new to the GDPR. Essentially, associations or industry bodies that represent a group of companies in an industry or sub-industry are encouraged to make a code of conduct for businesses they represent. Not many industries have done this, but you can find an example code of conduct by the cloud infrastructure services industry (available here: https://cispe.cloud/code-of-conduct/). Note that the association – not you – must create the code of conduct and have it approved by the Supervisory Authority.

How do I implement a Code of Conduct?

Inquire with the association or body representing your industry or sector regarding whether they have made a code of conduct. If so, your company then must sign to legally bind itself to complying with that code of conduct. It is also likely that the Code will require you to add additional details documenting the specific circumstances of your transfers and implementing the Code into your business operations.

Exceptions

There are a few exceptions outlined in Article 49 that allow you to make international transfers to third countries without needing one of the above legal mechanisms:

- **Data subject consent**: the person whose personal data you are transferring must explicitly consent to the transfer, after having been informed of the risks due to lack of legal safeguard.
- **Necessary for contract**: the transfer is necessary for the performance of a contract
- **Public interest**: it is in the public's interest to transfer data to a third country
- **Legal obligation**: the transfer is necessary for the establishment, exercise, or defense of a legal claim.
- **Vital interests**: the transfer is necessary to protect the vital interests of the data subject or of someone else, and the data subject is not physically or legally able to give consent.
- **Public register**: the transfer was made from a register that is intended to provide public information.
- **Limited circumstance**: if the transfer is not repetitive and concerns only a small amount of data subjects, and the controller has compelling legitimate interests that necessitate a transfer to a third country. (*Note that the controller must notify the Supervisory Authority of the transfer.)*