Can data privacy be achieved in blockchain?

Complying with database-driven regulations
in a blockchain world

Few technologies are sparking as much interest and excitement as blockchain. Concurrently, data privacy has emerged as a defining issue of our time, with more and more countries adopt laws granting new individual rights and protections regarding the processing of personal information. The problem? Many aspects of blockchain technology superficially appear to be incompatible with two of the most impactful modern privacy regimes, the European Union's General Data Protection Regulation (GDPR) and the California Consumer Privacy Act (CCPA). How can early blockchain adopters respond to privacy regulations today, while also preparing for future privacy developments? This article explores the intersection between blockchain and the most common shared domains of global privacy regulations—the collection, use, security, sharing, and deletion of personal information—with an emphasis on today's two most impactful privacy regulations, GDPR and CCPA.

How blockchain works

Breathless thought leaders have predicted blockchain will singlehandedly revolutionize industries as diverse as banking, healthcare, and identity management. Business verticals will be completely reimagined, as service provisioning and order fulfillment become decentralized and automated. But how exactly?

Blockchain is a decentralized, distributed ledger which incentivizes consensus-driven computation of hashes for historical veracity and integrity. As a decentralized system, blockchain lacks a single controlling entity, the way a large traditional database is administered by a single person or organization. A blockchain consists of a ledger, or historical record of transactions, which is distributed among all participants and is continuously compared against itself for accuracy. Individual transactions or records are stored in immutable blocks, which are hashed, and those hash values are used to compute subsequent blocks, forming a chain of dependent records.

If even one byte of information changes, the entire hash is altered regardless of the size of the block, as is the hash of each subsequent block. This provides an efficient means of checking for errors or alterations against other copies of the ledger without repeatedly accessing the information contained within blocks, which would require impractical amounts of time and computing resources. When a participant wants to add new information to the blockchain, a majority of participants must verify the new entry by performing a computational proof and arriving at a shared consensus that the information is accurate and consistent with prior blocks. Since computing proofs takes computational resources—and electricity, which someone has to pay for—participants may receive a small token of digital currency as incentive for performing validations or adding accurate information to the blockchain.

Blockchains can be public or private. Public, or "permissionless" blockchains allow anyone to access and contribute to the blockchain without authorization. By contrast, private or "permissioned" blockchains are only open to a defined group of individuals or organizations with access to the blockchain. Each uses a combination of public and private keys shared by individual participants to make transactions while ensuring the validity and integrity of those transactions. Both types of blockchains allow for more granular access controls. Owners of public blockchains can limit the right to contribute to certain participants, while owners of permissioned blockchains can restrict both the ability to view and to contribute to designated transactions within the blockchain.

How are blockchains used?

Blockchain-based systems are already in use today across a wide range of industries, with many more applications in development. BitCoin is the original blockchain use case, and the most well-known. As decentralized currency, BitCoin and other cryptocurrencies allow individuals and businesses to perform trusted transactions where traditional financial infrastructure is limited or nonexistent. The healthcare industry is in dire need of a system for managing patient health records, and blockchain-based solutions are under development which allow individuals to manage a single, secure health profile and share it with doctors or institutions as needed. By combining blockchain with the Internet-of-Things (IoT), logistics providers are developing automated supply chains which use smart contracts to reduce or remove the need for human intervention entirely. Realty, brokerage, and licensing industries are investing in blockchain technologies to develop systems for reliably tracking ownership and transfer of physical and intellectual property. While organizations should always perform a detailed assessment of whether blockchain is the most appropriate solution to a given business challenge, where blockchain provides an advantage, that advantage is substantial.

However, despite the innovative nature of blockchain applications, those applications face the exact same data privacy questions as their non-blockchain counterparts. What personal information is added to a blockchain, and who is responsible for it? How will that personal information be used? How is personal information shared, and with whom? How is the personal information secured? Finally, once the personal information serves its purpose and is no longer necessary, how will it be deleted? How organizations using blockchain will meet these privacy challenges deserves serious consideration.

Data collection

The first stage in any privacy analysis is identifying what information an organization collects, and how much of that information is personal information. In a typical blockchain setup, there are two forms of personal information for organizations to track:

- Personal information uploaded as a “payload” within blocks, such as customer records
- Personal information tied to blockchain miners or node operators, who must provide an IP address and other identifiable information for the blockchain to function.

Both GDPR and CCPA expressly identify IP addresses as personal information, so organizations running blockchains must treat transaction records and node data accordingly. However, these regulations also place the burden of managing collected information on the entity that owns the information—the “data controller” under GDPR, or the “business” under CCPA—which raises a larger question for organizations using blockchain. In a decentralized, distributed environment, who actually owns the personal information, and who is responsible for managing it?

Regulations like GDPR and CCPA place the burden of managing personal data on the entity that owns the information. But, for entities using blockchain, who actually owns such data—and who is responsible for managing it?

Establishing data ownership is crucial to navigating compliance with any privacy regulation, but poses a challenge for blockchain. Unlike centralized databases, no single entity controls and directs information added to a blockchain, except in some permissioned blockchain implementations with aggressive access controls. Without a single entity governing the information on the blockchain, users lack a clear point of contact for communicating individual rights requests, such as the otherwise simple request to access collected personal information. Without an effective means of receiving and responding to these requests, organizations relying on blockchain face an uphill battle in complying with modern privacy standards.

Addressing data ownership and access issues requires some forethought in how an organization implements its blockchain solution. The easiest way to manage personal information on a blockchain is to not store personal information on the blockchain at all. Where possible, information should be encrypted, hashed, or otherwise obfuscated so it is no longer identifiable without some form of private key controlled by the user. Personal information should be stored in a separate database. Even with these safeguards in place, organizations should still develop a clear, external-facing policy describing who is responsible for managing personal information on the blockchain and how individuals can contract that entity with individual rights requests.

The easiest way to manage personal information on a blockchain? Don’t store personal information on the blockchain at all.

Where collected personal information must be added to a blockchain, organizations should consider their justification, or “lawful basis” in GDPR terms, for doing so. Organizations relying on user consent must develop mechanisms for handling that personal information when consent is withdrawn. Most privacy regulations provide alternatives to consent-based collection and processing of personal information for limited business purposes, including GDPR and CCPA. As with traditional managed databases, organizations may also develop automated tools to manage individual rights requests in a decentralized capacity. A blockchain-based privacy dashboard, for example, would potentially allow users to access personal information stored about them on the blockchain, download a copy of that information, and even request modification of that information in a clear and simplified manner, without requiring knowledge of the underpinning technology or decentralized network of participants.

Use of data

While use cases for blockchain have received incredible attention over the last few years, the way organizations use personal information on a blockchain generally aligns with traditional personal information processing purposes. No matter how novel the mechanism, organizations are still processing information, which under both GDPR and CCPA encompasses nearly any possible activity an organization could perform on data it receives. These processes are subject to the same regulatory requirements as those occurring through traditional data storage models. However, there is one use case that is unique to blockchain, and that forms the basis for a substantial percentage of future blockchain solutions: smart contracts.

Smart contracts are small software programs loaded into a block that execute pre-programmed commands following some triggering event. The catch? Those algorithmic contracts can rise to the level of "automated decision-making" under GDPR, which raises privacy concerns and requires heightened protections due to the lack of human review or intervention. If a contract contains errors, either in the personal information it contains or within the contract's programming, those errors can only be corrected by nullifying the contract and adding a replacement contract on a new block—assuming the contract allows for those corrective measures in the first place.

To deal with these issues, organizations impacted by GDPR and executing smart contracts should ensure mechanisms are built into the original contract allowing for updates, pauses, and human intervention in response to complaints based on automated decision-making. To preempt these complaints before they arise, and to set appropriate expectations for users, organizations should ensure privacy notices are updated to reflect any smart contract processes and to provide a mechanism for filing complaints. Depending on the organization, external privacy policies also serve as an opportunity to explain novel technologies, such as blockchain and smart contracts.
This will:

- Help users understand how the technology works.
- Address privacy concerns before they arise.
- Build trust between the user and the organization.

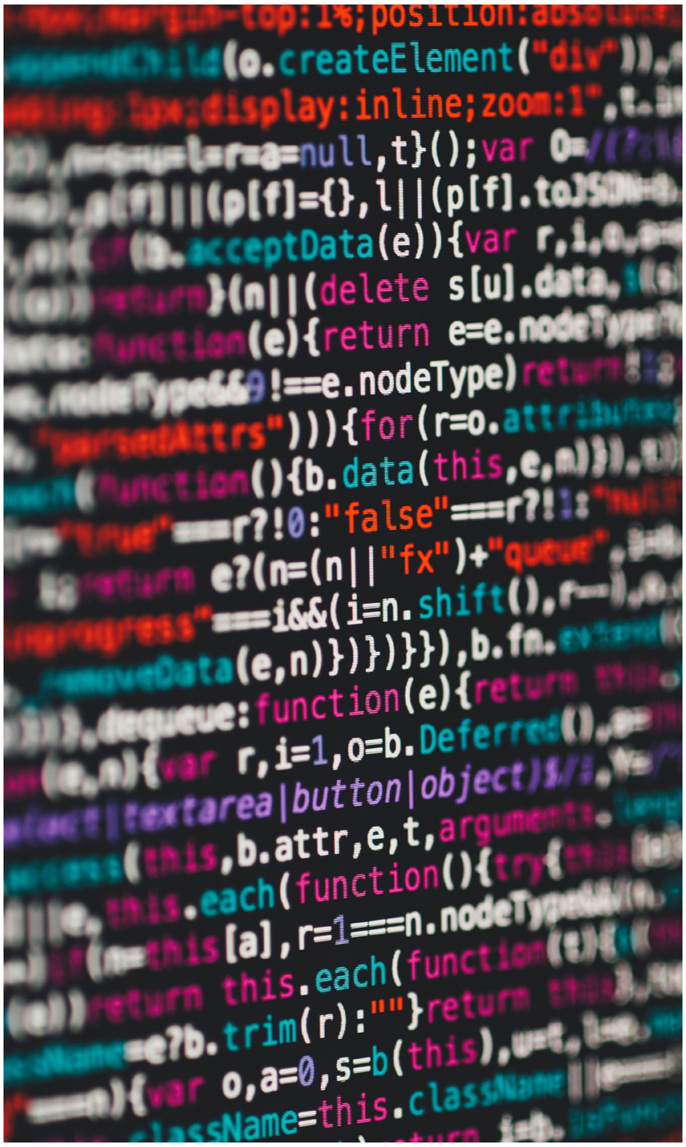
Where organizations must continue processing personal information stored in a smart contract for security, fraud prevention, record keeping or blockchain validation purposes, organizations should review the available lawful bases of processing under GDPR and determine where they have flexibility in responding to individual rights requests.

There is one use case that is unique to blockchain, and that forms the basis for a substantial percentage of future blockchain solutions: smart contracts.

Data security

Privacy regulations are generally technology-neutral, leaving it up to individual organizations or industries to determine reasonable and adequate levels of security. This allows a great deal of flexibility in implementing safeguards to protect blockchain systems. However, while blockchain offers unique solutions to traditional security threats, those solutions present unique vulnerabilities and do little to protect against modern social engineering attacks. Organizations managing data on a blockchain will still require a well-rounded and comprehensive security program, plus additional safeguards specific to blockchain.

To protect against rogue actors making unilateral, malicious changes to records, blockchain requires the consensus of a majority of participants to write information to the chain. However, if a group of participants—or a single participant with massive computing power—gains control of a majority of the nodes within a blockchain environment, that single participant or group may add new blocks or edit historical blocks within the record. While previously deemed a purely theoretical threat to blockchain, this so-called “51% attack” has led to several high-profile, multimillion dollar hacks for blockchain-based financial systems, with a notable uptick in just the last two years. Additionally, any business process is only as secure as its underlying software. The current blockchain ecosystem consists of a galaxy of third-party platforms, software interfaces, and custom iterations of contracts, code and currency. Each individual layer within a blockchain ecosystem carries the potential for software bugs and more traditional vulnerabilities due to hackers and delayed patches or updates. On top of this, threats associated with human error and social engineering remain. In fact, the majority of blockchain thefts occur through traditional phishing efforts, social engineering and insecure passwords.



Fortunately, building reasonable, appropriate, and compliant blockchain-related safeguards into an existing security program can be a relatively straightforward exercise. The following are four practices to strengthen blockchain security:

- Use permissioned blockchains with strong access controls instead of open, public blockchains, especially when storing personal information. Verifying participants and restricting access to critical chains can significantly reduce the threat of a 51% attack.
- Develop and enforce policies requiring timely and consistent updates to any equipment or software involved in maintaining a blockchain.
- Existing security training materials should be updated with blockchain-specific content, to further guard against phishing and related threats.
- Ensure that these measures are reviewed and updated at least annually to stay current with evolving threats, especially given the novel nature of many blockchain vulnerabilities.

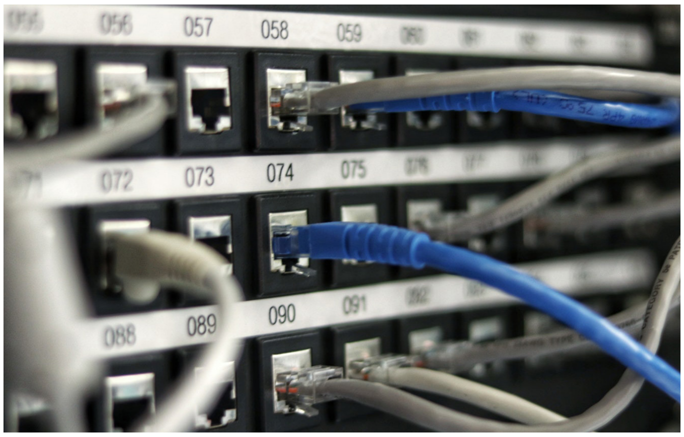
For organizations practicing privacy-by-design, those same principles can be applied to blockchain applications to ensure that, at the earliest stages of development, solutions are tailored so that personal data is minimized, adequately protected, and promptly removed once obsolete.

Data sharing

Around the globe, privacy regulations seek to limit the uncontrolled sharing of personal information with unknown third parties, but differing standards and requirements present a challenge to blockchain-based processes. For example, GDPR requires identification of a "data controller" and its downstream "processors"—effectively the entity which owns the personal information, and all its third-party service providers and sub-contractors doing something with that information on behalf of the data controller. GDPR places greater obligations and liability on data controllers, and requires contractual agreements with processors detailing how transferred information may be used. CCPA requires contractual agreements between businesses and service providers generally, but is less explicit on the contents of those agreements. CCPA also carries different obligations for personal information shared with service provides or sold to third parties. Yet neither statute addresses the core issue for blockchain implementations: what constitutes "sharing" on a distributed network where all participants have access to the entire transaction history by design?

Neither GDPR nor CCPA addresses the core issue for blockchain implementations: what constitutes "sharing" on a distributed network where all participations have access to the entire transaction history by design?

Under a strict interpretation of the term, an organization shares personal information with other entities whenever it adds a block containing personal information to a chain and distributes that information across all its nodes. Depending on applicable legal requirements, the organization may need to provide notice or receive consent from users before placing that information in a system accessed by other entities. However, where consent is required, users must be able to withdraw that consent, and the organization generally must communicate any opt-out decisions to entities with which it shared that personal information. Opt-outs raise similar issues as deletion requests, but require a more nuanced solution than simply masking personal information from all blockchain participants since opt-outs may be limited to certain processing purposes. Additional CCPA issues arise for businesses selling access to information stored on a blockchain.



While sharing and deletion issues are similar for blockchain-based systems, sharing issues generally require policy, not technical, solutions.

- Prior to using a blockchain to manage personal information, organizations should execute and update contractual agreements with related service providers, partners, and blockchain participants to reflect blockchain-specific concerns. Those contracts should clearly identify who owns or controls personal information added to the chain, define how that information may be used by participants, and state when and how participants must delete personal information they receive through the blockchain.
- Where possible, implement access controls to manage how information is shared among participants.
- Any organization adding personal information to a blockchain should review its privacy policy to ensure users receive sufficient notice of how personal information is used and how to contact the data owner, both for compliance purposes and to set appropriate expectations for users.
- Lastly, organizations should establish a mechanism for confirming shared personal information is not being misused or retained longer than permitted, to avoid a Facebook-style Cambridge Analytica scenario. More than ever, organizations are expected to police their own data sharing agreements, and may face increased liability for failing to do so.

Data deletion

Both GDPR and CCPA provide individuals with the right to request that organizations delete their personal information. These rights are not absolute—there are many exceptions, which organizations should leverage to triage and reduce the volume of deletion requests they receive. Additionally, once an organization flags personal information for deletion, it still retains flexibility in determining how to delete that information. In most cases, rendering personal information non-identifiable is sufficient, meaning anonymization and some forms of pseudonymization are just as permissible as a hard delete. This can allow the organization to comply with the deletion request without losing the business intelligence tied to that user's account. These processes are relatively straightforward in today's structured databases, but require foresight when applied to blockchain implementations.

By design, blocks cannot be altered once added to a chain. Since hashed values of blocks are used to compute the hash of subsequent blocks, changing or deleting a historical block value requires modification of all blocks added after it. Those changes would require approval from a majority of nodes and significant computing resources to recalculate and verify subsequent hashes, making individual deletion requests non-scalable. It's the equivalent of convening a board meeting to update an email address.

The simplest solution to these issues is to not store personal information on a blockchain. Multiple cryptographic methods are available today that allow organizations to store hashed values or commits within a blockchain. Those values typically refer back to personal information stored in a separate, traditional database. This raises separate issues regarding interoperability between traditional databases and blockchains, but the demand for blockchain-based record keeping ensures these issues will be addressed with future technology.

Where it is absolutely critical to keep personal information on a blockchain, certain statutory exceptions mentioned above could help justify continued processing of personal information despite individual requests. For example, under GDPR, an organization may rely on contractual agreements with users or legitimate business interests as justifications for not deleting personal information, provided the proper agreements and risk assessments are performed. Under CCPA as currently written, businesses may argue the continued processing of personal information is necessary as an internal business use to maintain the integrity of the blockchain, provided that use aligns with customer expectations. In either case, organizations will have to justify that the continued processing of personal information deserves to overrule the individual's right to have personal information deleted. Organizations should anticipate these regulatory challenges when implementing any blockchain-based record management systems, and develop blockchain-related compliance and policy strategies early on to minimize organizational impact.

> While some aspects of blockchain may seem incompatible with certain privacy obligations, these issues are not insurmountable.

Modern privacy regulations are not mere compliance exercises. They establish genuine rights to data privacy for individuals around the globe and will only continue to expand over the next several decades. In the rush to innovate new blockchain-driven solutions for today's problems, developers and organizations must ensure they are doing so while honoring the rights of their customers, users and clients. Generally, organizations can and should take the following steps to reduce privacy risks when adopting blockchain:

- Perform a thorough review to ensure blockchain is the most appropriate solution to a given use case.
- Don't store personal information on a blockchain, or where absolutely necessary, ensure procedures are in place to permanently de-identify stored personal data.
- Update contracts and service agreements to address access, ownership, and terms related to blockchain-based systems.
- Update internal and externally-faced policies to address to adequately prepare employees and establish user expectations for the blockchain-based service.

Though some aspects of blockchain may seem incompatible with certain privacy obligations, these issues are not insurmountable. If blockchain is going to be, as many think, the dawn of a new era, that new era can and will include adequate protections for individual rights to data privacy. For now, complying with privacy regulations while using blockchain is a tricky needle to thread, but with the proper guidance, insights, and innovations, organizations can achieve these goals while controlling privacy risks and honoring individual rights to data privacy.

Grant Thornton

GT.COM