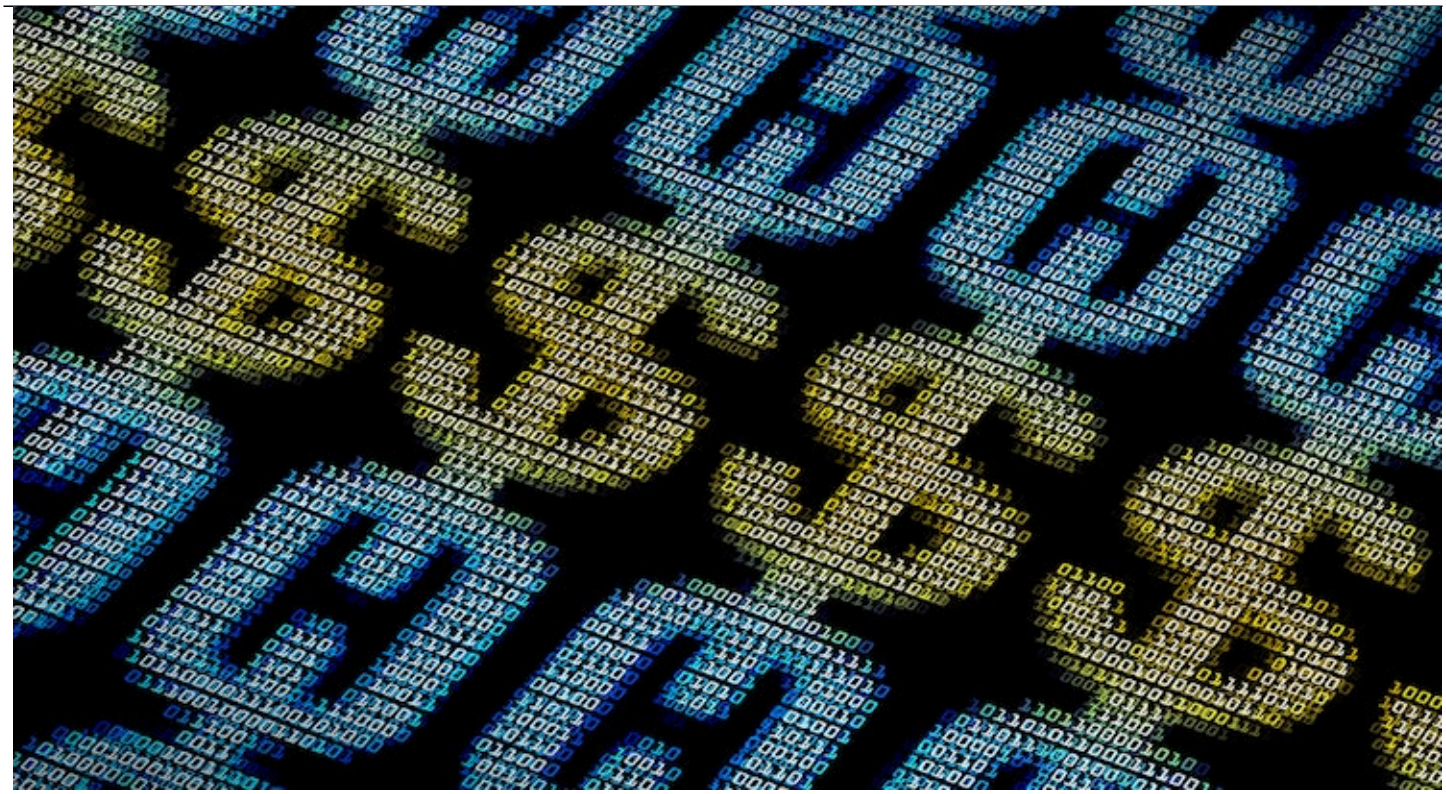
ACC DOCKET
INFORMED. INDISPENSABLE. IN-HOUSE.

Blockchain Technology is Here — Is it Compliant with GDPR?

Technology, Privacy, and eCommerce



CHEAT SHEET

- ***Controller.*** Under the EU's General Data Protection Regulation (GDPR), a data controller determines the purpose and means of processing personal data. The controller in a permissioned blockchain is either the single entity or consortium that owns it. In a permissionless blockchain, it's unclear if it is the developer, participating nodes, or validating nodes.
- ***Erasure.*** GDPR asserts an individual's right to be deleted, but blockchain is an immutable database unless it is made redactable, which is more plausible in a permissioned blockchain.
- ***Jurisdiction.*** Permissionless blockchains stretch across jurisdictions, but permissioned blockchains can include location restrictions to reduce cross-border data transfers.
- ***Compatibility.*** At this time, permissioned blockchains are the most compatible with GDPR.

There have been many articles discussing blockchain technology over the preceding few years. And nearly as many addressing Europe's General Data Protection Regulation (GDPR). This article asks a fundamental question: Are blockchain innovations compatible with GDPR? The opinions range from inherent conflict to nearly perfect harmony. However, most of them agree that there is no uniform definition of "blockchain technology." Instead, there are many applications implemented within a set of blockchain technologies, and each of them must be assessed on its own.

GDPR, as its name implies, protects data. Blockchain technology is so compelling because it protects data through decentralization, transparency, and immutability.

This article argues that some blockchains present a risk of incompatibility with GDPR, whereas others can be used in a GDPR-compliant manner. The tensions between GDPR and blockchain revolve around three areas: (1) the anonymization of personal data; (2) identification and obligations of data controllers and processors; and (3) data subjects' rights.

Remind me: What is a blockchain?

A blockchain is a system of distributed ledgers hosted on a group of connected computers that together agrees on the status of, and changes to, shared data. As a rule, blockchain applications use cryptographic technologies.

There are different types of blockchains. Each can be classified by how access to transaction data is granted:

1. In a public blockchain reading access and a right to execute transactions are granted to every user and node. The most famous example is Bitcoin.
2. In a private blockchain reading access (i.e., the capability of auditing the blockchain) and a right to execute transactions are granted to a predefined list of users and nodes. An example is RippleNet, a global payments network.
3. In a permissionless blockchain, writing access (i.e., capability of validating transactions and adding a block as well as voting in a consensus mechanism) are granted to every node (i.e., every node can verify transactions and add new block). An example is Holochain, a distributed computing platform.
4. In a permissioned blockchain writing access is granted to a predefined list of nodes. An example is Corda, an open-source blockchain project that allows businesses to create smart contracts with relevant parties.

GDPR and blockchain

GDPR, as its name implies, protects data. Blockchain technology is so compelling because it protects data through decentralization, transparency, and immutability.

GDPR defines personal data as any information relating to an identified or identifiable natural person.

Personal data includes pseudonymous data but excludes anonymous data.

Under GDPR, "pseudonymization" means the processing of personal data so that it can no longer be attributed to a specific data subject without the use of additional information. The additional information must be kept separately and should be subject to technical and organizational measures to ensure that it's not possible to attribute it to a natural person.

In contrast, anonymous data is data that cannot be attributed to a natural person either by itself or with additional information.

So, is data processed on a blockchain pseudonymous or anonymous? And does this data fall within

the scope of GDPR?

Blockchain protocols employ public-private-key cryptography, also known as asymmetric cryptography, and cryptographic hash functions.

Asymmetric cryptography

Asymmetric cryptography is used to identify users of a blockchain (e.g., through accounts) and to authorize transactions. Account numbers are open to the public. The owner of an account must use a private key — which is not known to the public — to execute transactions relating to that account. For example, a Bitcoin public address looks like this: 37aEHh9ME3kU7AZ3rUxBCyKR5FhR3hbqVn.

> One analytics platform, CrystalTM, took only three hours to track payments in bitcoins from the victims of the WannaCry virus

It is impossible to identify a natural person with this string of numbers and letters only. But, with some time and resources, it is possible to do this indirectly with the help of additional information. For example, users may reveal their public key if they are asking for a donation or if they are selling or buying goods or services. Some blockchain platforms can also identify individuals by examining transactions. One analytics platform, CrystalTM, took only three hours to track payments in bitcoins from the victims of the WannaCry virus. Additionally, attackers may use IP addresses, cookies, and other metadata to link a public address with an individual. For these reasons, blockchains with public addresses should be considered pseudonymous data.

Hash functions

Every transaction on a blockchain uses hash functions.

Hash functions take data of any size and transform them into a unique string of numbers and letters. The phrase “Hello world!” generates this hash value: C0535E4BE2B79FFD93291305436BF889314E4A3FAEC05ECFFCBB7DF31AD9E51A. If the nearly 20,000 words of GDPR are entered, the hash value is the same size: 0B6CA03E54DEA013928D81AC9A35F0092A6AEC63CA27DF7F2A23FC8C5A0745F2.

Hash values are unique. Any small change greatly alters an output hash value. For example, if we delete a space in the title of GDPR by changing it from “REGULATION (EU) 2016/679…” to “REGULATION(EU) 2016/679…” the previous hash value will change completely: 2231DA0ED36A1171A8A3FFAAA085150701913DBA3267EB74F525082921D70B18.

Thus, hash values are digital fingerprints of data. Changes are easily detectable.

In addition, hash functions are one-way functions. It is extremely difficult (but not impossible) to reverse-engineer hash values into input data. So, are hash values of personal data anonymous or pseudonymous data?

On the one hand, a risk remains that a brute force attack, which systematically tries millions of combinations, could reverse a hash value if there are a limited number of possibilities for the data. For example, if the hash value contained the number of employees in a small or even a medium company, and the attacker knew it could only be a certain number, the attacker can easily compute

all the possible values of X to find the outcome. While there are obfuscation techniques, the fact that it's possible to reverse engineer the hash value is the important point. If it's possible, then it means that a hash value of personal data is not anonymous but pseudonymous data, which is covered by GDPR.

Who is the data controller on a blockchain?

GDPR defines a "data controller" as any natural or legal person who determines the purpose and means of the processing of personal data. The controller is the main person who is responsible for compliance with GDPR.

Identifying a controller on permissioned blockchains is straightforward. Permissioned blockchains are run either by a single entity or by a consortium of entities. They determine the purpose and means of data processing on permissioned blockchains. Thus, either a single entity in a private blockchain or all or some members of the consortium should be treated as joint controllers.

However, GDPR was written with the centralized data management model in mind. A permissionless blockchain is decentralized so identifying a controller on a permissionless public blockchain is tricky. There are three types of entities that may qualify as a controller on a permissionless public blockchain: developers, participating nodes, and validating nodes.

Developers of blockchain protocols craft the architecture of blockchain systems — they do not determine the purpose and means of actually processing personal data. Therefore, they are unlikely to be treated as controllers.

A node on a blockchain is a device that maintains the blockchain. Each participating node can be treated as a data controller if its transactions are professional or commercial activities. The French Data Protection Authority (CNIL) agrees with this definition. Its "Blockchain: Premiers éléments d'analyse de la CNIL," gives the following example: If a notary records a transaction on behalf of a client, the notary is a controller because it determines the purpose (e.g., change of ownership) and means of processing of personal data (data format, blockchain application). Participating nodes, whose activities cannot be qualified as professional or commercial activities (e.g., users of a Bitcoin blockchain transacting virtual currency on their own behalf) should fall under the "household" exemption of GDPR, which states the regulation does not apply to personal data processing when it is done "by a natural person in the course of a purely personal or household activity."

Validating nodes simply run an algorithm embedded in a blockchain and, thus, might not be treated as determining the purpose and means of processing of personal data.

> If your business is considering blockchain products, it is wise to steer them to permissioned blockchains until more questions can be answered about permissionless blockchains.

Permissionless blockchains raise many questions when it comes to data ownership. It is up to the courts, data protection authorities, and the legal and technology community to determine whether they will be compatible with GDPR. If your business is considering blockchain products, it is wise to steer them to permissioned blockchains until more questions can be answered about permissionless blockchains.

Rights and obligations under GDPR and blockchains

GDPR provides the following rights to individuals:

1. The right to be informed
2. The right of access
3. The right to rectification
4. The right to erasure
5. The right to restrict processing
6. The right to data portability
7. The right to object
8. Rights in relation to automated decision-making and profiling

Like with the data control question, the real antagonism between GDPR and blockchain technologies stands out when a permissionless public blockchain is used.

For example, which node or nodes should obtain consent of a data subject when no controller can be found? How should such consent be obtained, given that it must be specific and unambiguous?

If it is unclear who is the controller on a permissionless public blockchain, then whom can the data subject request the right of access from (i.e., a right to know what data are processed, for what purpose, who the data are shared with, and so on)?

A permissionless public blockchain is distributed across multiple geographical locations. These locations are neither predefined nor fixed. Therefore, jurisdictional and cross-border data transfer issues arise and become overly complicated to deal with on permissionless public blockchains.

One reason blockchains are popular is that they are very difficult to edit. To rewrite data, the entire structure needs to be rewritten — not only where the slight edit needs to be made. The accumulated costs of rewriting (e.g., due to recalculating hash puzzles) make it extremely hard to manipulate the transaction history on a permissionless public blockchain. As a result, the blockchain becomes an immutable database. This is the core nonfunctional feature that ensures integrity in a purely distributed peer-to-peer system.

Assuming the controller is identified, how can the GDPR’s right to erasure be realized, given this immutability feature? What constitutes erasure in blockchain systems? Assuming that personal data is encrypted, destroying the key renders the data unreadable. Is this enough to comply with GDPR if the data is technically still there?

There are no definitive solutions to these issues right now. How to deal with these issues is an ongoing debate. However, it is possible to limit the risk of incompatibility with GDPR in implementing blockchain technologies.

Making blockchain more compatible

Given the earlier points, a permissionless public blockchain runs a high risk of being incompatible with GDPR. The root of this conflict lies in a purely distributed peer-to-peer network architecture and its immutability. To make a blockchain more compatible you need to choose a “permissioned” blockchain.

First, permissioned blockchains are run either by a single entity or by a consortium of entities. A single organization has the authority to add nodes and also grant them particular reading and/or writing rights on a private blockchain. In a consortium blockchain, all members, a set of members, or a single member grants reading and/or writing rights.

They determine the purpose and means of processing of data on blockchains. They are well known or easily identifiable. Therefore, the issue of controllership, either the sole or joint version of it, is easy to deal with in the case of permissioned blockchains.

Second, since the controller(s) decides which nodes can be admitted to the network, they can also add restrictions with regard to the locations of those nodes. Therefore, the issue of cross-border data transfer can also be dealt with in permissioned blockchains fairly easy.

Third is the “immutability” issue. This is the most difficult one to deal with from legal and technical points of view.

One approach is to make a “redactable” blockchain.

In this type of blockchain, a lock is added to each link of the hash chain. This function contains a special key to a lock — known as a trapdoor. Without the knowledge of the trapdoor, it is hard to open the lock and to redact the block. Thus, the chain remains immutable like a typical blockchain.

But, with the knowledge of the trapdoor, it is possible to replace the content of any block. Moreover, it is also possible to delete, modify, or insert any number of blocks. Furthermore, if the trapdoor key is lost or destroyed, then a redactable blockchain reverts to an immutable one.

In the case of a purely private permissioned blockchain, a trapdoor key could be given to a central authority so that, under certain circumstances, it can redact the blocks.

In the case of a consortium permissioned blockchain, a trapdoor key can be shared by all the parties to the consortium, and redactions can be realized using a secure multiparty computation protocol.

Another approach is the implementation of an off-chain data storage architecture. Under this approach, personal data are stored off-chain in a centralized or distributed manner and hash references to that data are processed on a blockchain.

When a data subject requests the erasure of certain data, then a data processor can delete a link between a blockchain hash reference, a pointer, and the off-chain data. However, such a deletion provides a “logical,” rather than purely technical, erasure. Depending on the risk of re-creation of the deleted link, such an approach may be compatible with the concept of erasure.

A GDPR-compatible blockchain use case

At Asters, we are testing blockchain with an off-chain data storage architecture. This private permissioned blockchain is based on the Emercoin platform.

Data subjects input the personal data through web interface secured with HTTPS. The back-end system encrypts the data, computes hash IDs, puts encrypted data with hash IDs onto storage, and inserts hash IDs into the Emercoin name-value storage. Each transaction also has a unique transaction hash ID. So, the data hash IDs are stored on the blockchain. The encrypted data is stored

off-chain. Access to data is given through a special key for decryption. The blockchain can keep records of data transactions for audit purposes because each manipulation has its own hash ID.

The data can be changed and appropriate hash IDs can be updated due to the unique functions of the Emercoin name-value storage. Deletion can be done as soon as a retention period expires. Deletion is also possible when requested by the data subject.

The above architecture and protocol ensure compatibility of this blockchain-based solution with GDPR. The controller can be easily identified; the data subject's consent can be obtained in a specific and unambiguous manner; the right to know what data is processed, for what purpose, and who the data was shared with can be enforced; writing and reading rights can be granted in a compatible way; data are processed in a secured manner; and the data subject can control personal data, including requesting its erasure.

Conclusion

In general, GDPR is a technology-agnostic regulation. It is not hard to see that most provisions of GDPR have been written with a centrally governed database in mind. Accordingly, purely distributed databases based on peer-to-peer networks, so-called public permissionless blockchains, may conflict with GDPR.

Surely, there may be use cases where the design of a particular blockchain system can smooth out the frictions between certain properties inherent in the blockchain and GDPR. Moreover, blockchain is not a static concept that, once developed, remains unchanged. It is subject to constant developments undergoing not only technical improvements but also conceptual advancements.

Thus, each use of blockchain should be carefully analyzed on its merits to come to a conclusion about its compatibility with the requirements of GDPR. However, to be on the safe side, private permissioned blockchains are the best-suited model to process personal data.

ACC EXTRAS ON… Blockchain

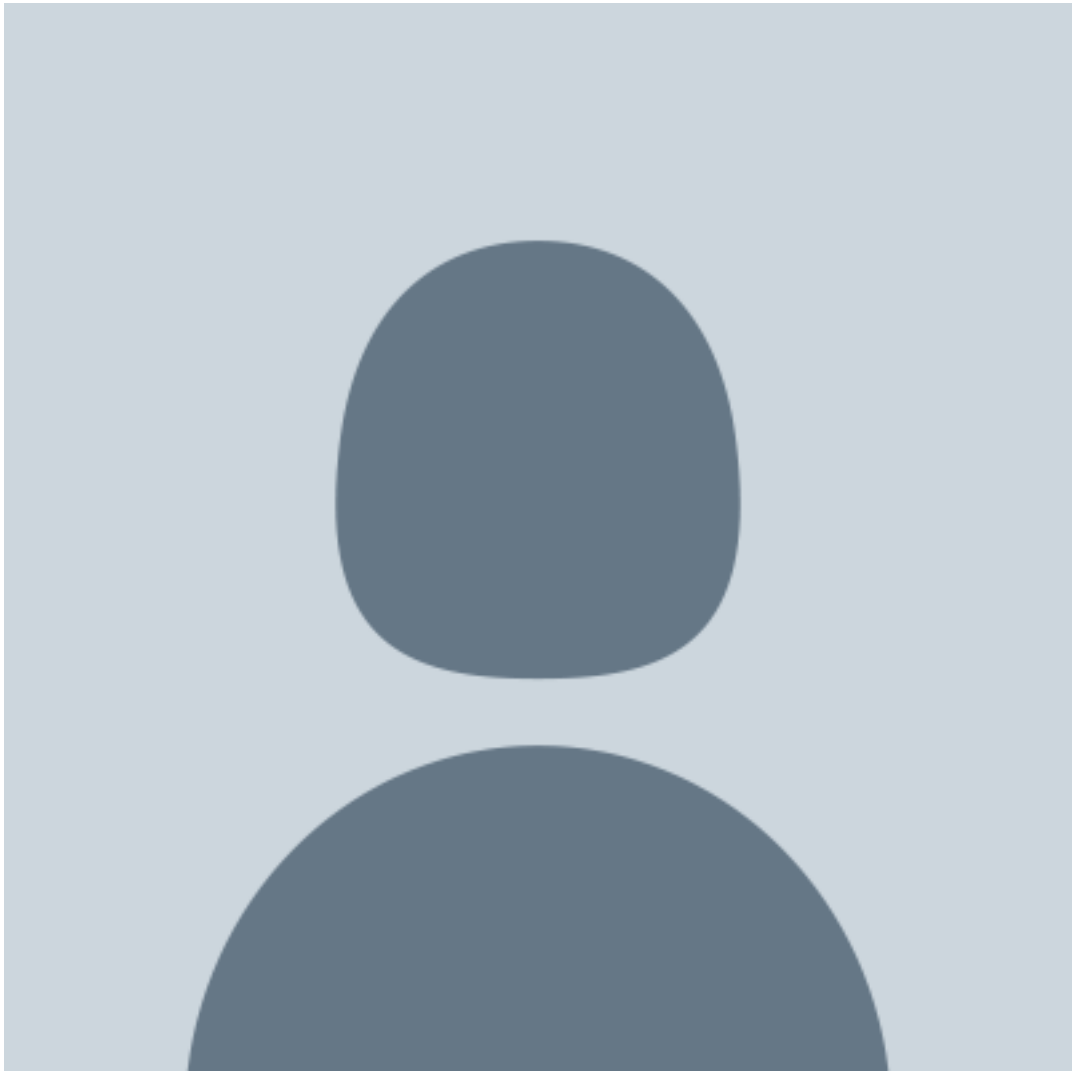
ACC Docket

Blockchain Basics: Blockchain in Action (Dec. 2019).

Blockchain Basics: Global Regulations (Nov. 2019).

Blockchain Basics: What GCs Need to Know About the Disruptive Technology (Nov. 2019).

Elene Dighmelashvili

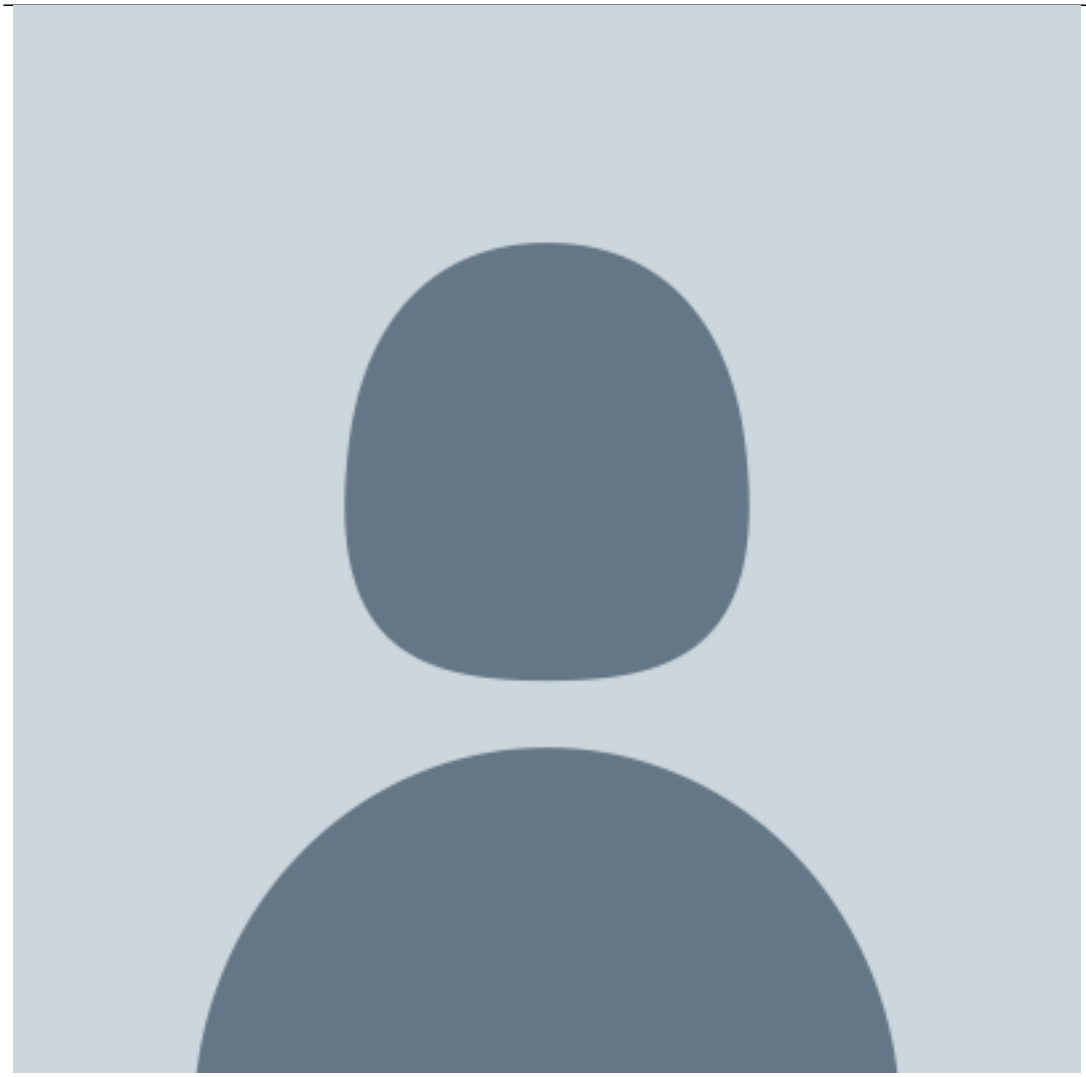


Legal Counsel

BitFury Group

Elene Dighmelashvili is legal counsel for BitFury Group, a blockchain technology company.

Aleko Nanadze



Counsel

BitFury Group

Yuriy Kotliarov

Partner

Asters

Yuriy Kotliarov is a partner at Asters, a Lex Mundi member firm for Ukraine, and former general counsel at Ukrteleco.

Sergiy Tsyba

Counsel

Asters

Sergiy Tsyba is counsel at Asters, a Lex Mundi member firm for Ukraine.