

# **BLOCKCHAIN AND GDPR**

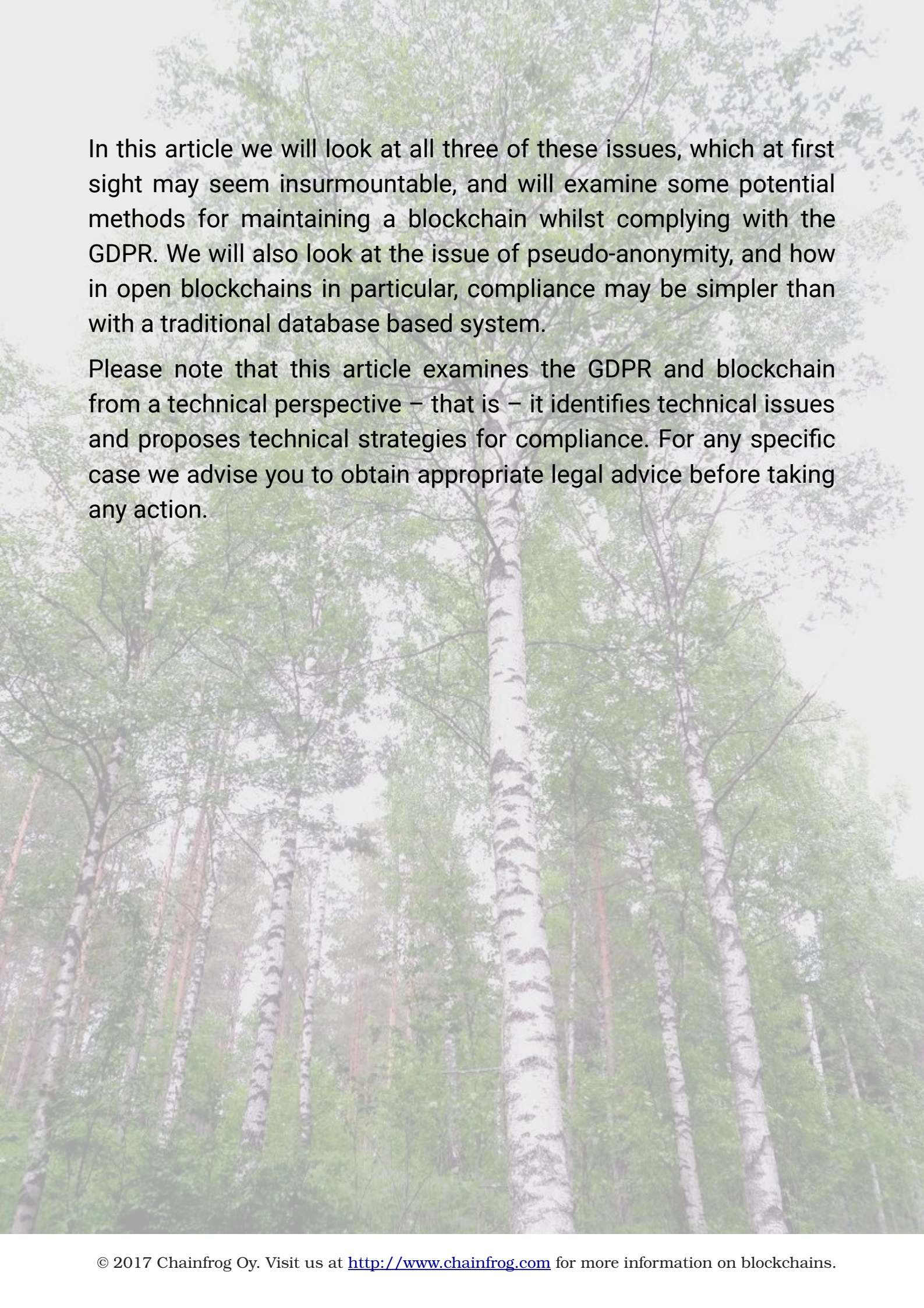
# BLOCKCHAIN AND GDPR

## HOW TO SQUARE PRIVACY AND DISTRIBUTED LEDGERS

In 2016 the European Union passed the General Data Protection Regulation (GDPR) in order to give European residents more rights and control over their personal data. It comes in to full force on 25 May, 2018, and will affect any company holding data relating to private EU citizens or residents, whether or not the company holding the data is based in Europe. Compliance will be essential, as penalties can be as much as the higher of 4% of worldwide turnover or 20 million euros.

The regulation has profound significance to blockchain systems in three regards:

- Data stored on a blockchain is tamper proof, so deleting it later on is not an option.
- Blockchains are distributed, so control of the data put on them is relinquished.
- Smart contracts will fall under the auspices of automated decision-making, and may therefore be contested.



In this article we will look at all three of these issues, which at first sight may seem insurmountable, and will examine some potential methods for maintaining a blockchain whilst complying with the GDPR. We will also look at the issue of pseudo-anonymity, and how in open blockchains in particular, compliance may be simpler than with a traditional database based system.

Please note that this article examines the GDPR and blockchain from a technical perspective – that is – it identifies technical issues and proposes technical strategies for compliance. For any specific case we advise you to obtain appropriate legal advice before taking any action.



# A BLOCKCHAIN PRIMER

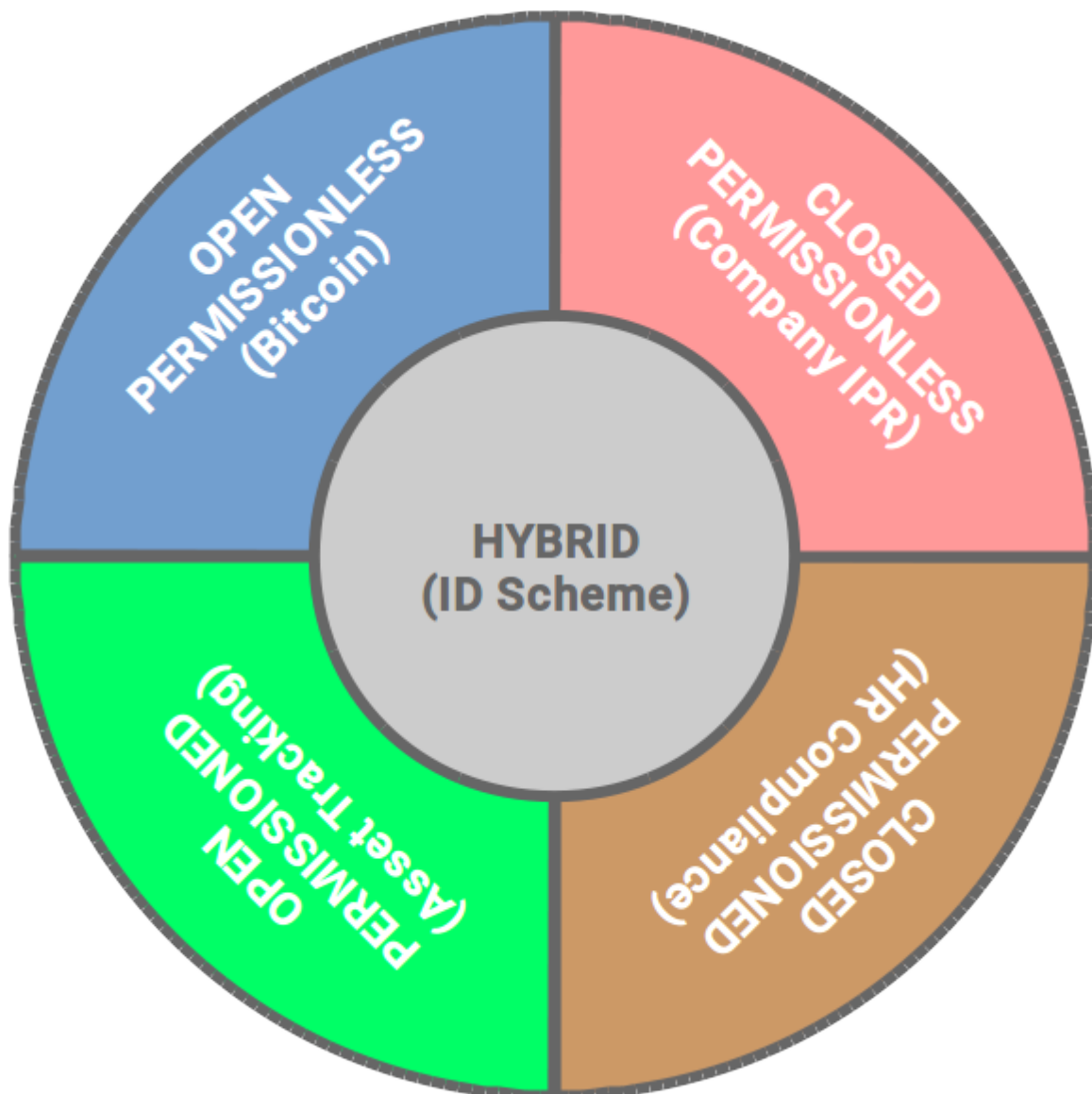
Let's start by examining some of the terminology and properties of blockchains that are significant, and have a brief look at their technical underpinnings.

## **Public and private, open and permissioned**

In order to grasp the issues faced by blockchains that will not affect traditional databases, it is important to be aware of the different kinds of blockchains that have been proposed over the last few years. The terms that need to be understood are: public versus private, and open versus permissioned.

The first blockchain system was Bitcoin, and as the system was designed to allow anyone with a computer to submit transactions or join in with maintaining the network, it is both public (Bitcoin runs on the internet), and open (anyone can create a bitcoin address, or download or design software to run nodes that perpetuate the Bitcoin network).

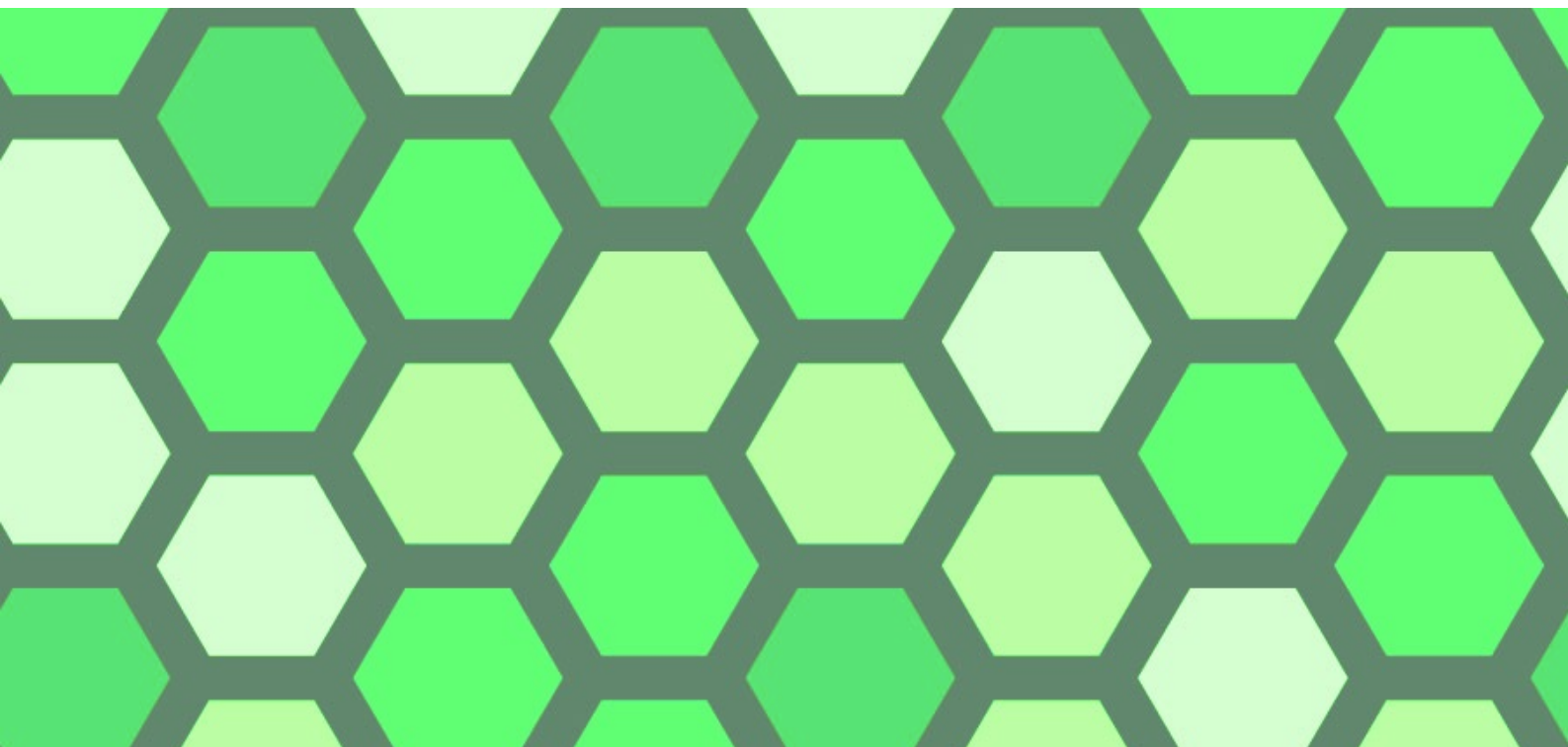
At the other extreme, we see private and permissioned blockchains. For these systems, the blockchain is run on a private network, for example a VPN or intranet, and an administrator has to grant permission to any individual wanting to submit transactions or maintain a blockchain node. Such a blockchain might, for example, be used by an HR department within a large corporation, where there is a need to provide an auditable record of HR data, but the company does not want just any employee to view or add to the blockchain data, and it certainly doesn't want the public to see the blockchain data content.



Completing the basic possible combinations, we have open and permissioned blockchains (anyone can see the blockchain, but only specific agents can add data; a use case could be a distributed social media system or a digital rights management system), and private yet permissionless blockchains (for example, a corporate whistleblowing blockchain, in which any employee can submit a complaint or notification, but people outside of the company can't see the data recorded on the blockchain, and company officers cannot identify the reporting person).

Finally, there may also be hybrid systems, for example an open blockchain in which anyone can submit a transaction, but only specific permissioned computers are allowed to generate blocks and maintain the blockchain system.

The difficulties that are faced in complying with the GDPR increase the further from a traditional private database system your blockchain configuration is. A private permissioned blockchain is almost the same as a private database system, whereas a public open blockchain is a very different beast indeed, and a whole new set of compliance problems arise.

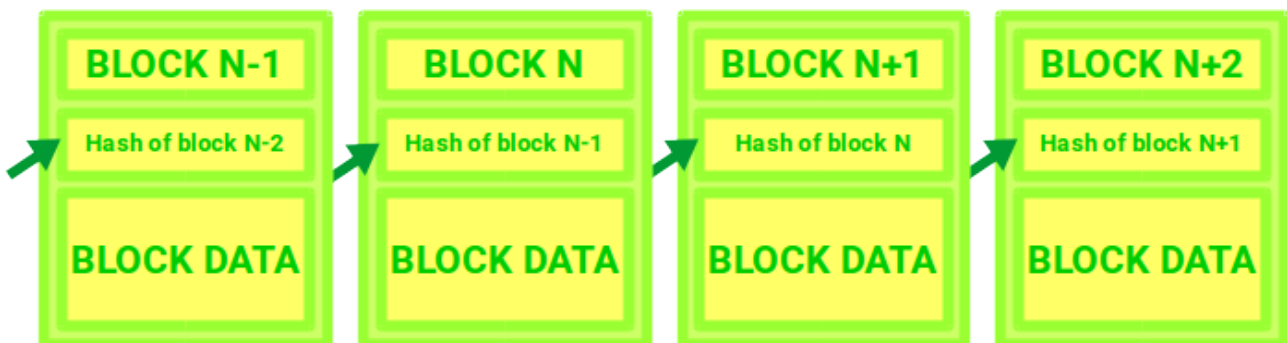


## Tamper proof and distributed

One of the great benefits that blockchains bring, is tamper proof recording of data. This gives blockchains the ability to record financial transactions or other data in a manner that allows subsequent auditing of the data for veracity and authenticity, and ensure that, for example, when a transfer of ownership or value is recorded, the transaction cannot subsequently be reverted. In essence, it is the tamper proof nature of blockchains that allows them to create cryptocurrencies, by making unique unforgeable digital entities.

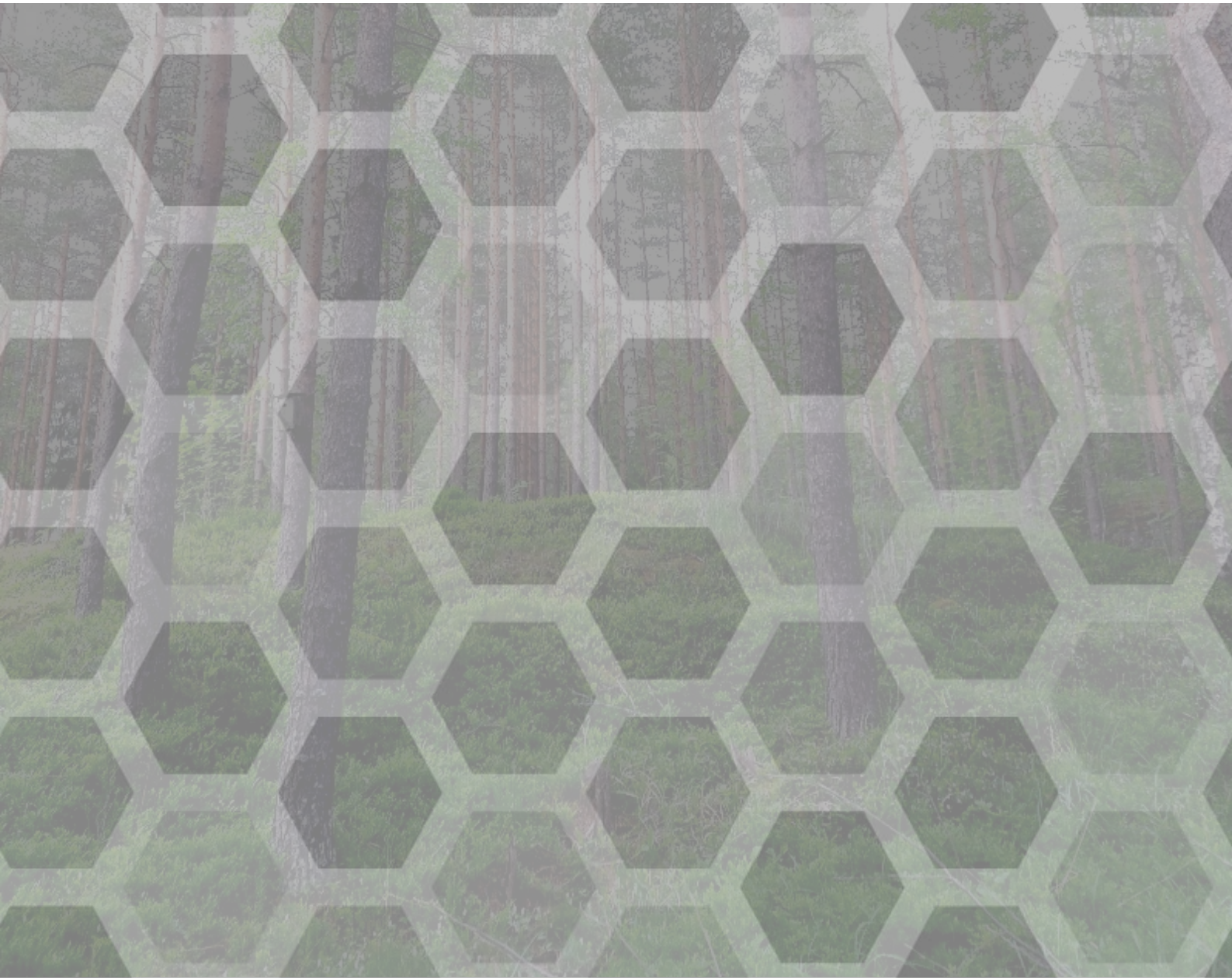
As they are generated, the records or data are submitted by clients to the peer-to-peer network running the blockchain, and nodes on the network package the unprocessed data into blocks, which are added to the existing chain at regular intervals. Thus the chain grows over time.

To ensure that the chain of blocks cannot be tampered with, each block contains a reference to the preceding block by including a cryptographic hash of the data within the preceding block. If the data in a block is altered, the hash of the block changes too, and this falsification of the records can therefore be detected.



Furthermore, because a complete record of the chain of blocks is kept by all nodes maintaining the blockchain, if an individual node has a copy of the data that has been tampered with, not only can this be detected, but a valid copy can be requested from one of the other nodes. The node with falsified data may even be blacklisted by the remaining nodes and expelled from the blockchain system.

Because the data on the blockchain cannot be deleted or modified at a later date, at first sight it appears that it is not possible comply with some of the edicts of the GDPR, in particular the “right to erasure”, whereby an EU resident may request that their details be deleted from the system.





## **Contesting smart contracts**

A smart contract is a snippet of code that is stored on a blockchain, and which is run by the nodes maintaining the blockchain when the conditions specified in the code are met. As a result a smart contract can be used in a transparent and automated way to, for example, hold funds in escrow and release them when the required contract terms are met, or to issue tickets for an event and automatically handle refunds if the event is cancelled. See our eBook on smart contracts to understand how they work, and what limitations and risks they currently carry.

Over time, it is expected that smart contracts will evolve and develop to make automated decisions that could be based on EU residents' data. As a result, the decisions made and actions taken can be contested under the GDPR legislation.

If the smart contract is out there "in the wild" on a public blockchain, and it doesn't contain any code to deal with the reversal of decisions it makes, then it won't comply with GDPR, and the issuer of the smart contract could be held liable in perpetuity. But if the creator or owner of the contract cannot be identified, this issue becomes moot.

# HOW TO RESOLVE THE CONFLICT

Having detailed the technical properties and capabilities of blockchains that are relevant to the GDPR, we now turn our attention to possible solutions to the problem of reconciling the GDPR edicts with the blockchain system.

## **The right to be forgotten**

Although there are a substantial number of conditions under which private data may still be retained, for example for compliance with superseding legal obligations, when the data is in the public interest, or when freedom of expression takes precedence, in general EU residents may request that their data be transferred to another data storage provider, or even deleted entirely. This clearly presents a problem to a blockchain system, in which data cannot be deleted. There are number of solutions:

## ***Do not record personal data on a blockchain***

The most obvious method to sidestep the GDPR is simply not to put any individual data on the blockchain relating to any private citizen or resident of the EU. However, this drastically reduces the usefulness of blockchains for any public application, such as health record tracking, social media, reputation reporting systems associated with online sales, and identity systems such as an international passport. The GDPR does not specify if subsequent corrections to the data are acceptable, if the original incorrect data is still present in earlier blocks on the blockchain.

### ***Record personal data pseudo-anonymously***

Blockchains allow data to be associated with individuals pseudo-anonymously. For example, in Bitcoin it is possible to see which bitcoin addresses have balances, and from which other addresses the bitcoins were transferred, but it is not possible without a forensic investigation to determine the identity of the person who controls the bitcoin address. Similarly, provided the data being recorded on the blockchain does not identify the individual, a system can be put in place whereby, for example, dietary preferences, hobby interests, purchases and so on can be recorded and linked via a pseudo-anonymous address. However, postal addresses, phone numbers, and even IP addresses cannot be recorded using this method as they can be used to track down the person behind the data.

### ***Encrypt the data on the blockchain***

A further possibility is to ensure that all private data stored on the blockchain is encrypted. In such a situation, the company responsible for data care can provide evidence of the deletion of the data by ensuring that the decryption key is destroyed. Another approach may be to shift the responsibility for protecting the private key to the individual whose data is being stored on the blockchain. However, this is a risky strategy, because if the key is leaked, the data is no longer protected, and cannot be removed. This leads us to our current preferred solution:

### ***Store the data in a referenced encrypted database***

Another approach is to store the relevant data on a private encrypted database, and include a hash of the data on the blockchain. The hash can be used to confirm that the data in the database has not been tampered with, but no actual identifiable data is present on the blockchain itself.

Copies of the decryption key for the data may also be stored on the blockchain, with each key copy encrypted with the public key of the various agents that are allowed to access the data. If there is a requirement to delete data, the relevant records or tables in the database can be dropped. If access needs to be changed or restricted, the data can be decrypted and then re-encrypted with new keys. Interestingly, this also provides a method for the transfer of data between controllers, without having to physically move it.

### **Contesting automated decisions**

Given that the ruling grants EU persons the right to contest automated decisions, and smart contracts running on a blockchain are effectively making automated decisions, the GDPR needs to be taken in to account when developing and deploying smart contracts that use personal data in the decision making process, and produce a legal effect or other similarly significant effect.

### ***Smart contract over-rides***

The simplest means of ensuring smart contract compliance is to include code within the contract that allows a contract owner to reverse any transaction conducted. There are however a number of problems that could arise from this. For example, if subsequent transactions have been conducted based on the original decision, all these need to be rolled back too. As the appeal time can be long, many such actions may have been taken after the original contract decision, and it may not even be possible to roll back all the actions.

### ***Consent and contractual law***

A second approach is to ensure that the users activating the smart contract are aware that they are entering into such a contract, and that they provide explicit consent. The GDPR provides the possibility of waiving the contesting of automated decisions under such terms,

but the smart contract would require putting on hold any subsequent actions to be taken until consent is obtained.

### ***Decentralized apps and the GDPR***

It's not all plain sailing on the other side either. What if a decentralized application (Dapp) is launched anonymously on a public blockchain, for example on Ethereum, which gathers information on citizens, collects and redistributes funds, and makes automated decisions based on the data and funds available? According to the EU ruling, users of the Dapp have the right to have their data expunged from the blockchain, and to contest decisions made by the Dapp. But in a true Dapp there is no board of directors or CEO to hold accountable, no bank account that can be frozen, and no data controller to petition. Banning the relevant blockchain and enforcing the ban by putting pressure on internet service providers and VPN suppliers is unlikely to be effective. As the old adage goes – the internet treats censorship as a malfunction and simply routes around it.



## SUMMARY

Blockchains are already being used to store data pertaining to individuals, for example in digital rights management, and in the future the amount of personal data on them will just grow and grow. If blockchains are to be used to track customer purchases, allow consumers to trade energy between different electricity providers, prove ownership of cars or assets, or any other applications that require personal data, then the organizations developing and launching such systems will have to abide by the GDPR.

At the simplest level, the requirement to allow former users of the system to “be forgotten” is the most pertinent, and we have described a number of technical approaches that allow a blockchain system to benefit from the tamper proof and auditability properties of the system, while still allowing the sensitive data to be transferred between controllers or even deleted.

We have also examined the more complicated area of smart contracts, and in particular the right to contest automated decision making.

As is always the case, additional legislation brings additional overheads, and the GDPR in relation to blockchain is no exception. However, as we have shown, there are plenty of solutions waiting to be tried out, and as we gain more clarity on how the EU intends to apply the GDPR in a practical manner, so too the approaches to blockchain implementation and application should become clearer.

# REFERENCES

EU GDPR Portal: <http://www.eugdpr.org/eugdpr.org.html>

GDPR legislation markup site: <https://www.privacy-regulation.eu/en/>

UK ICO – getting ready for GDPR: <https://ico.org.uk/for-organisations/resources-and-support/data-protection-self-assessment/getting-ready-for-the-gdpr/>

Decentralizing Privacy: Using Blockchain to Protect Personal Data;  
Zyskind, G., Nathan, O. and Pentland, A. :  
<http://web.media.mit.edu/~guyzys/data/ZNP15.pdf>