IT-Security and Usability

Chinmay Khandekar, Daniel Theis

1 INTRODUCTION

IT-Security and blockchain are both very challenging topics since they are novel and constantly changing due to their nature. But, on the other hand, the blockchain is used in our everyday life to complex systems. In this report, we have highlighted the Security and Data management challenges when implementing blockchain systems. IT-Security is a comprehensive risk management system with technical aspects, cybersecurity frameworks, integrity assurance, and best security practices to reduce the impact of risk against cyber-attacks and fraud. Therefore we narrow down onto the data protection aspects of GDPR along with the TCDP, C5, and other BSI Grundschutz standards.

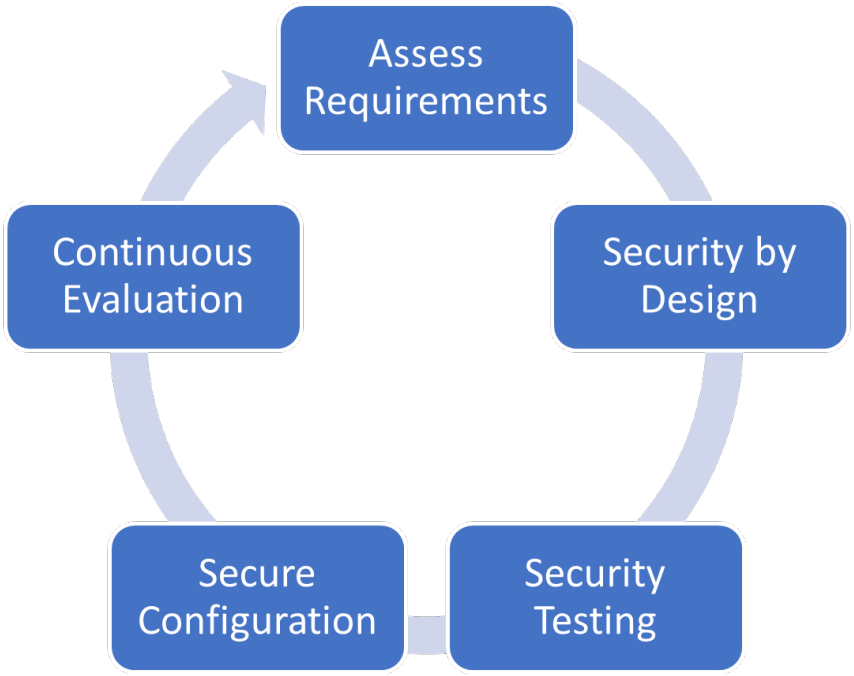
The report consists of several parts in which we highlight some security considerations on designing GDPR and TCDP compliant systems. For example, we come across techniques for managing PII (Personally Identifiable Information) in the information flow under the GDPR regime with pseudonymization (2.2.1) and anonymization (2.2.2) strategies available during this report. Going on further in Section 3 we discuss some vital overlapping requirements in the TCDP and GDPR. The report then summarizes some key user security considerations when interacting with blockchain in light of increasing crypto-currency wallet theft.

2 SECURITY CONCEPT

The foundation basis of IT-Security is the CIA Triad which is confidentiality, integrity, and availability. In contrast, the concepts relating to the people who use that information are authentication, authorization, and non-repudiation. Examples include research data, personal data, and confidential technical configuration. In this part of the report, we have listed some security concepts to consider when designing and developing a system.

2.1 Security Assessment

The concept of a Security concept is to manage IT-Security risks and comply with standards and best practices to secure and ensure the system's compliance. To ensure the system is "secured," we have to have some security baseline or assessment against which it can be evaluated. To make the process continuous, we have a simplified version to ensure security and compliance. In order to come up with a concise security concepts we refer to various globally accepted standards such as NIST 800-37 [21], ISO 31000[20] and how CDC [4] manages Cyber Risk and how IT-Security and Data Protection and referred to the GDPR [14], TCDP[23] and BSI C5 [18].



1. Assess Requirements: It is important to assess the level of protection for the system at the first stage where clear distinction on the data to be protected is defined [15]. Once the level of protection is determined then, appropriate data security measures can be formulated based on the requirements:
    - Permissions: The data can be accessed or altered or read or deleted only by those you have authorization to do so (the users that have the permission do so with the right intent;
    - Data Processing Purpose: The person or the organization has complete relation is to why it is processing the data.
    - Data Security: The data held must be backed up to prevent accidental data loss, alteration, or destruction by ransomware or other threat agents. Therefore ensuring data security and backups form a critical requirement in regards to data.
2. Security by Design: Articles 25(1) and 25(2) of the GDPR outline the obligations to

enforce data protection by design and default [14]. In comparison, this can be one way of demonstrating compliance with these requirements. When considering data protection by design, it has a broad application [7]. The idea is to keep security in mind. Some examples to consider security by design when:

- Implementing a new blockchain node
- The physical security of equipment which is used for identification (i.e., card readers)
- inter-application processing and flow of data or
- using personal data for another purpose than the defined purpose.

3. Security Testing: The idea of security is to enforce and test the critical aspect to ensure the security control mechanisms are in place and functioning as defined in the requirements. Where it is essential to ensure the appropriate functioning with the intended results, some examples include:
   - Barcode scanners are displaying personal information publicly instead of pseudonymizing personal details.
   - The users can access the confidential page of the portal without authentication.
4. Secure Configuration: It is essential to apply appropriate security to the data in the form of encryption and cryptographic algorithm to protect personal data. These are some considerations:
   - When implementing encryption, it is important to consider four things: choosing the suitable algorithm, choosing the correct key size, choosing the appropriate key vault, and keeping the key secure.
   - It is critical to ensure the right key algorithm and key size to protect against any attack or brute force.
   - The encryption software plays a critical part. The encryption must meet current standards such as FIPS 140-2 and FIPS 197 [9].
   - Ensure that the private keys are kept secure and have processes to generate new keys in case of a compromise.

   5. Continuous Evaluation: IT security is dynamic and constantly evolving, with new attack vectors, and vulnerabilities discovered almost every day. Therefore continuous evaluations form a critical step to ensure IT-Security effectiveness with the importance to evaluate compliance. We have listed some examples you can include in the checklist:
   - Controls: Keeping the user permission and access-list updated for various parts of the network and nodes.
   - User Security: Ensuring appropriate and sufficient access rights are assigned to the user to carry out their duty.
   - Security Configuration: It is essential to ensure the security configuration is set to Read-only, and only appropriate users can only edit these.

- Password Policies: Passwords are essential in the access authentication of IT-Systems. Therefore, it is important to set a complex password requirement policy with strength complexity and expiry for force changing. Key-Pairs still form as a more secure way to authenticating compared to passwords for critical systems.
- Vulnerabilities: a secure software can be insecure due to a vulnerability discovered, such as in the encryption algorithm of keys that can eventually make them vulnerable and insecure. Therefore regular assessment on whether the encryption method remains appropriate should be undertaken.
- Updates: Ensuring systems and interconnected systems are regularly to ensure critical vulnerabilities and 0-days in the system is patched.

2.2 Privacy

The era of big data analytics promises varied data sources. It offers opportunities to extract hidden values from unstructured raw data sets through novel reuse. However, the reuse of personal data is a crucial concern for data protection law. It involves processing for purposes beyond those that justified its original collection, at odds with the principle of purpose limitation [19].
The processing of personal data attracts the attention of data processing laws and where user information is involved privacy plays a greater role in its protection. While processing such personal data it is possible by either pseudonymize or anonymize the personal data aspect. The significant difference between the two concepts relates to the goals to achieve with these techniques [8]. The discussion two methods are discussed in their parts below.

2.2.1 Pseudonymization

Pseudonymization aims at protecting personal data by concealing the personally identifiable information in the data, by replacing either one or more personal data identifiers with pseudonyms, and covering the link between the aliases and the original identifiers. An identifier is a specific text of information holding a privileged and close relationship with an individual, which may allow direct or indirect identification of this individual (e.g., name, email address, location, could be a picture of an individual, MAC address of hardware, IP address, etc.) [6, 16].

The critical aspect is that pseudonymization separates the original dataset into two parts, where each of the parts has meaning regarding specific individuals only in combination with the other [6, 16]. There are several pseudonymization techniques with some examples:

- Hashing without key: A cryptographic hash function h is a function with specific properties which transforms any input message m of arbitrary length to a fixed-size output h(m) (e.g., of size 256 bits, that is 32 characters), being called hash value or message digest [6, 16].
- Hashing with key: A robust approach to generate pseudonyms based on the use of keyed hash functions – i.e., hash functions whose output depends not only on the

input but on a secret key too; in cryptography, such primitives are being called message authentication codes. The controller shall keep the private key securely stored separately from other data. It constitutes the additional information, i.e., it provides the means for associating the individuals such as the original identifiers – with the derived pseudonyms [6, 16].

- Hashing with key and salt: the input to the hash function is being augmented via adding auxiliary random-looking data that are being called "salt" [6, 16].
- Encryption symmetric or asymmetric (i.e., usage of public keys) aims to ensure via proper use of mathematical techniques that the whole dataset that is being encrypted is incomprehensible to anyone but authorized users who are allowed to reverse/decrypt the dataset. To this end, encryption is the main instrument to protect personal confidentiality of data by hiding the whole dataset and making it unintelligible to any unauthorized party (as long as state-of-the-art algorithms and key lengths are used and the encryption key is appropriately protected) [6, 16].
- Tokenisation: it refers to the process that the data subjects' identifiers are replaced by randomly generated values, known as tokens, without having any mathematical relationship with the original identifiers. Hence, knowledge of a token has no usefulness for a third party other than the controller or processor [6, 16].
- Other well-known techniques, such as masking, scrambling, and blurring. All of them mainly focus on pseudonymizing data being at rest (i.e., data that is being stored in a file or database) [6, 16].

2.2.2 Anonymization

Anonymization can be defined as a method in which the information can be manipulated (concealed or completely hidden) to make it nearly difficult to directly identify the data subjects [13, 12]. There exist some anonymization techniques with some examples:

- Randomization: anonymization techniques that can alter the accuracy of the data to remove any strong link between the data and the individual. The technique induces some random addition, permutation, or combinations with a differential privacy mindset.
- Generalization: this is an anonymization technique that generalizes, or weakens, the attributes of data subjects by modifying with a greater magnitude. This family includes techniques such as aggregation or K-anonymity, L-diversity, and T-closeness.

3 Discussion

In this section we will discuss from the user point of view in regards to data requirements in the purview of TCDP and GDPR..We have come up with data essentials in respect to TCDP and GDPR:

Requirements	Compliance
Data Controller DPO.	GDPR Article 25, Articles 38 and 39.
Internal Access Controls	TCDP No. 22, ISO/IEC 27002 point 11.1, GDPR Recital 78
Data Processing Rights	TCDP No. 23, GDPR Article 6.
Encryption and Data Privacy	TCDP No. 11, GDPR Article 32
Deletion, Correction and Blocking Data Access	TCDP No. 6, ISO IEC 27018 point A.1.1, GDPR Article 15 and 16.
Data related Disclosures	TCDP No. 8, ISO/IEC 27018-point A.5.2, GDPR Article 33.
Data Deletion or Right to be Forgotten	TCDP No. 10, ISO/IEC 27018 point A.9.3, GDPR Article 17
Data Requests and Access or Right of access by the data subject	TCDP No. 9, ISO/IEC 27018 point A.9.3, GDPR Article 20
Secure Data Exchange	GDPR Article 5
Data Policy	GDPR Article 12

3.1 BLOCKCHAIN-GDPR

The GDPR, and more broadly classical data protection principles, were designed in a world in which data management is centralised within specific entities. In this respect, the decentralised data governance model used by blockchain technology and the multitude of actors involved in the processing of data lead to a more complex definition of their role. Not all actors involved in the blockchain are data processors. Miners are only validating transactions submitted by participants and are not involved in the object of these transactions: therefore, they do not define the purposes and the means of the processing. The Smart contract developers who process personal data on behalf of the data controller are within the means of GDPR [3]. Blockchain poses several challenges around GDPR with multiple open-ended questions:

- What is considered personal data on a blockchain? Wallet address? Transaction? Hashed personal data?
- Who, in a distributed environment, is the data controller, data processor, or even joint controller?
- How to deal with the right to be forgotten (GDPR Art. 17) on an immutable ledger?
- Who is to be held responsible in case of, for example, a breach?
- Can be encrypted personal data be stored on a blockchain?

Firstly, the GDPR is based on an underlying assumption that concerning each personal data point, there is at least one natural or legal person – the data controller whom data subjects can address to enforce their rights under EU data protection law. These data controllers must comply with the GDPR's obligations. Blockchains, however, are distributed databases that often seek to achieve decentralization by replacing a unitary actor with many different

players. The lack of consensus as to how (joint-)controllership ought to be defined hampers the allocation of responsibility and accountability [1].

Second, the GDPR assumes that data can be modified or erased where necessary to comply with legal requirements, such as Articles 16 and 17 GDPR. However, blockchains render the unilateral modification of data purposefully onerous to ensure data integrity and increase trust in the network. Furthermore, blockchains face the challenges of adhering to data minimization requirements with the purpose limitation in the current form of the data economy [1].

When implementing blockchain, there has to be a definitive architecture to keep in mind regarding the separation between personal and non-personal data to apply GDPR or data protection law. Therefore the concept of pseudonymization is one of the novelties of the GDPR [1].

In the blockchain context, public keys serve as the kind of identifiers mentioned in Recital 30 GDPR. Blockchains rely on a two-step verification process with asymmetric encryption. Every user has a public key (a string of letters and numbers representing the user), best thought of as an account number shared with others to enable transactions. In addition, each user holds a private key (also a string of letters and numbers), which is best thought of as a password that must never be shared with others. Both keys have a mathematical relationship by which the private key can decrypt data that has been encrypted through the public key. Public keys thus hide the identity of the individual unless they are linked to additional identifiers. This is only the case where the public key relates to a natural person. There are DLT use cases where public keys do not relate to natural persons. For example, where financial institutions use a blockchain to settle end-of-day inter-bank payments for their own accounts, public keys would relate to these institutions and not natural persons, meaning that they would not qualify as personal data subject to the GDPR [1, 11].

A public key is data that 'can no longer be attributed to a specific data subject' unless it is matched with 'additional information' such as a name, an address, or other identifying information, and thus pseudonymous data according to Article 4(5) GDPR. [10] Indeed, there are many analogies between public keys and other pseudonymous strings of letters and numbers, such as unique identifiers in cookies, which have been said to qualify as personal data. Beyond, public keys may also reveal a pattern of transactions with publicly known addresses that could 'be used to single out an individual user', such as through transaction graph analysis [2, 1].

3.1.1 Private and Permissionless Blockchains

In private or permissionless DLT, there is generally a determined legal entity (such as a company or a consortium) that determines the means and, in many cases, also the purposes of personal data processing. Where this is the case, that entity qualifies as the data controller. However, there may also be joint controllers in such circumstances. In line with Wirtschaftsakademie Schleswig Holstein, it can be argued that those using such infrastructure for their

purposes are joint controllers. An example would be a consortium blockchain established between many actors in the same supply chain. The legal entity created by the consortium would be a controller considering that it exercises significant control over the purposes and the means of personal data processing. Yet, the individual companies that have joined the consortium and are subsequently using the infrastructure for their purposes, thus enabling the DLT to process new personal data, could also qualify as joint-controllers [1].

3.2 User Security

There is an emerging trend beyond cryptocurrency payments: the blockchain could enable a new breed of decentralized applications without intermediaries and serve as the foundation for key elements of Internet security infrastructures [22]. Often, we read the news of stealing money from exchanges, servers, and cryptocurrency owners. A big challenge in Bitcoin and all blockchain cryptocurrencies is securing private keys. Blockchain uses elliptic-curve asymmetric cryptography to control the ownership of coins or accounts. For example, to transfer a coin from a user to another, the sender user signs a transaction with her private key. The blockchain network verifies the transaction's signature with its public key. After being verified and accepted by the blockchain network, the transaction, unlike the traditional bank transfer, cannot be rolled back by anyone [17]. The wallets have been primary targets for hackers due to the inadequate security in the cryptocurrency portals, smart contracts, or tokens. In recent times based on the security incidents and wallet related thefts, these are some security considerations for users to implement to safeguard their blockchain wallets [5].

- The users should enable 2-step authentication for their wallets. These can be OAuth 2.0 based authentication such as Google Authenticator or SMS-based authentication for web-based transactions. The 2-step authentication prevents misuse of unknown transactions from your wallet.
- Encrypting your wallet: Creating complex passwords for wallets and writing down the password somewhere safe is an important aspect to consider, given the limited password recovery options for the wallet.
- Using a strong password: Any password that contains only letters or recognizable words can be considered very weak and easy to break. Passwords must be 16 characters long with alphanumeric or password manager generated complex passwords. More complex passwords are hard to brute force.
- Private Keys. The private keys are the only identification mechanism to access the wallet remotely. Therefore storing the private key secure and enabling authentication over the web should be enforced with utmost precaution.
- Hardware Wallets are now alternative to software-based wallets for the transaction. They are more secure since they are hardware-based and not remotely accessible over the internet unless connected to the PC/Laptop.

4 SUMMARY

The concepts developed are in line with the GDPR application on blockchain and IT-Security. In comparison, it is easier to comply with GDPR with a Private Permissioned Blockchain due to the highly centralized operations with clearly linked identifiers of individuals to ensure the right to forgetting and the right to erasure.

Our report has highlighted some best IT-Security and Blockchain practices considering Data protection law and security standards. Blockchain is a technology that will bridge multiple sectors and implementations in finance, logistics, healthcare, and cybersecurity. At the same time, the report has predominately focussed on GDPR and TCDP. Some research works in the Blockchain and IT-Security context. Still, since the field is constantly developing, there is a possibility to research blockchain applications and cybersecurity aspects since new malicious actors and attack vectors are created daily.

Future research opportunities in IoT-Blockchain-Cyber Security also exist as the world moves towards adopting encryption, distributed IOTs, and increasing forms of distributed monitoring securely with cryptography, encryption, and security. Alternate research opportunity exists in the cyber-security aspect of fungible cryptocurrency tokens. The increasing usage of cryptocurrency and tokens have attracted unwanted attention by cybercriminals for illicit activities such as ransomware and terrorism financing. Therefore, the Forensics and Law enforcement investigation into the trial could also be a suggested future research topic.

REFERENCES

[1] *Blockchain and the General Data Protection Regulation*. en. July 2019. URL: https://www.europarl.europa.eu/RegData/etudes/STUD/2019/634445/EPRS_STU(2019)634445_EN.pdf.

[2] F. Z. Borgesius. "Singling out people without knowing their names - Behavioural targeting, pseudonymous data, and the new Data Protection Regulation". In: *Comput. Law Secur. Rev.* 32 (2016), pp. 256–271.

[3] CNIL. *Solutions for a responsible use of the blockchain in the context of personal data*. en. 2018. URL: https://www.cnil.fr/sites/default/files/atoms/files/blockchain_en.pdf.

[4] Centers for Disease Control and Prevention. *Vulnerability Management Life Cycle*. en. 2021. URL: https://www.cdc.gov/cancer/npcr/tools/security/vmlc.htm.

[5] Michael Fröhlich, Felix Gutjahr, and Florian Alt. "Dont lose your coin! Investigating Security Practices of Cryptocurrency Users". In: *Proceedings of the 2020 ACM Designing Interactive Systems Conference* (2020). DOI: 10.1145/3357236.3395535.

[6] Joshua Gresham. *Is encrypted data personal data under the GDPR?* en. Jan. 2018. URL: https://iapp.org/news/a/is-encrypted-data-personal-data-under-the-gdpr/.

[7] *Guide to the UK General Data Protection Regulation (UK GDPR).* en. July 2019. URL: https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/.

[8] Mike Hintze and Khaled El Emam. "Comparing the benefits of pseudonymisation and anonymisation under the GDPR". In: *Journal of Data Protection Privacy* 2.2 (2018), pp. 145–158. ISSN: 2398-1679. URL: https://www.ingentaconnect.com/content/hsp/jdpp/2018/00000002/00000002/art00005.

[9] *Implementation Guidance for FIPS 140-2 and the Cryptographic Module Validation Program.* en. Mar. 2003. URL: https://csrc.nist.gov/csrc/media/projects/cryptographic-module-validation-program/documents/fips140-2/fips1402ig.pdf.

[10] Data Protection Commission Ireland. *Guidance on Anonymisation and Pseudonymisation.* en. 2019. URL: https://www.dataprotection.ie/sites/default/files/uploads/2019-06/190614%5C%20Anonymisation%5C%20and%5C%20Pseudonymisation.pdf.

[11] Christopher Millard Jatinder Singh Jean Bacon Johan David Michels. "Blockchain Demystified: A Technical and Legal Introduction to Distributed and Centralised Ledgers". In: *25 Rich. J.L. Tech.* 1 (2018).

[12] Jasmien César Julien Debussche. *Big Data Issues Opportunities: Anonymisation Pseudonymisation.* en. 2019. URL: https://www.enisa.europa.eu/publications/privacy-and-data-protection-by-design.

[13] Paul Ohm. *Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization.* July 2012. URL: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1450006.

[14] European Parliament and the Council. *Regulation (EU) 2016/679.* 2016. URL: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A02016R0679-20160504&qid=1532348683434.

[15] *Privacy and Data Protection by Design.* en. Jan. 2015. URL: https://www.enisa.europa.eu/publications/privacy-and-data-protection-by-design.

[16] *Recommendations on shaping technology according to GDPR provisions - An overview on data pseudonymisation.* en. Jan. 2018. URL: https://www.enisa.europa.eu/publications/recommendations-on-shaping-technology-according-to-gdpr-provisions/at_download/fullReport.

[17] Hossein Rezaeighaleh and Cliff C. Zou. "New Secure Approach to Backup Cryptocurrency Wallets". In: *2019 IEEE Global Communications Conference (GLOBECOM).* Waikoloa, HI, USA: IEEE, Dec. 2019, pp. 1–6. ISBN: 978-1-72810-962-6. DOI: 10.1109/GLOBECOM38437.2019.9014007. URL: https://ieeexplore.ieee.org/document/9014007/ (visited on 07/24/2021).

[18] Bundesamt für Sicherheit in der Informationstechnik. *Kriterienkatalog C5*. en. 2021. URL: https://www.bsi.bund.de/EN/Topics/CloudComputing/Compliance_Criteria_Catalogue/C5_NewRelease/C5_NewRelease_node.html;jsessionid=7481BDC80F513BEB8AB34443F2B3F59B.internet082.

[19] Sophie Stalla-Bourdillon and Alison Knight. "Anonymous Data v. Personal Data — A False Debate: An EU Perspective on Anonymization, Pseudonymization and Personal Data". In: 2017.

[20] International Organization for Standardization. *ISO 31000*. en. URL: https://www.iso.org/iso-31000-risk-management.html.

[21] National Institute of Standards and Technology. *NIST SP 800-37 Rev. 2*. en. 2018. URL: https://csrc.nist.gov/publications/detail/sp/800-37/rev-2/final.

[22] Paul J. Taylor et al. "A systematic literature review of blockchain cyber security". en. In: *Digital Communications and Networks* 6.2 (May 2020), pp. 147–156. ISSN: 23528648. DOI: 10.1016/j.dcan.2019.01.005. URL: https://linkinghub.elsevier.com/retrieve/pii/S2352864818301536 (visited on 07/24/2021).

[23] *Trusted Cloud – Data Protection Profile for Cloud Services (TCDP) – Version 1.0*. en. Sept. 2016. URL: https://tcdp.de/data/pdf/14_TCDP_v1.0_EN.pdf.